Modellierung von Angebots- und Nachfrageverhalten zur Analyse von Agrarpolitiken: Theorie, Methoden und empirische Anwendungen

Henningsen, Arne

Department of Agricultural Economics, University of Kiel

August 2006

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

Modellierung von Angebots- und Nachfrageverhalten zur Analyse von Agrarpolitiken: Theorie, Methoden und empirische Anwendungen

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

vorgelegt von
Dipl.-Ing. agr. Arne Henningsen
aus Flensburg

Kiel, im Mai 2006

Dekan:	Prof. Dr. Joachim Krieter
Erster Berichterstatter:	Prof. Dr. Dr. Christian H.C.A. Henning
Zweiter Berichterstatter:	Prof. Dr. Awudu Abdulai
Dritter Berichterstatter:	Prof. Dr. Thomas Heckelei
Tag der mündlichen Prüfung:	20. Juli 2006

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

Die Arbeit kann als pdf-Dokument unter
http://e-diss.uni-kiel.de/agrar-fak.html
aus dem Internet geladen werden.

Danksagung

Mit der Drucklegung dieser Dissertation endet ein wichtiger Abschnitt meines Lebens. Nach Abschluss dieser arbeitsreichen, aber sehr schönen Zeit möchte ich denjenigen Menschen danken, die entscheidend zur Fertigstellung dieser Arbeit beigetragen haben.

An erster Stelle möchte ich meinem Doktorvater Herrn Prof. Dr. Dr. Christian H.C.A. Henning für die engagierte Betreuung dieser Arbeit danken. Auch durch die mir gewährten Freiräume hat er wesentlich zum Erfolg der Arbeit beigetragen. Herrn Prof. Dr. Awudu Abdulai und Herrn Prof. Dr. Thomas Heckelei danke ich für die Übernahme des Zweit- bzw. Drittgutachtens.

Insbesondere bedanke ich mich bei meinen Co-Autoren Prof. Dr. Dr. Christian H.C.A. Henning, Dr. Alison Burrell, Jeff D. Hamann, Carsten Struve, Jörg Müller-Scheeßel und Anja Wiebusch für die angenehme und konstruktive Zusammenarbeit. Bei Prof. Dr. Dr. h.c. Ulrich Koester und Dr. Alison Burrell bedanke ich mich dafür, dass sie mein starkes Interesse an der agrarökonomischen Forschung geweckt und somit den Weg für diese Arbeit bereitet haben. Weiterhin danke ich den vielen freiwilligen Entwicklern der verwendeten freien Software wie z.B. R, LaTeX, Subversion, KDE und GNU/Linux.

Wesentlich zum Gelingen der Arbeit trug auch das angenehme Arbeitsumfeld und die freundliche Atmosphäre am Lehrstuhl für Agrarpolitik und den Instituten für Agrarökonomie sowie Ernährungswirtschaft und Verbrauchslehre bei. Dafür danke ich meinen Kolleginnen und Kollegen. Für die gewissenhafte Durchsicht der Beiträge und für viele hilfreiche Verbesserungsvorschläge danke ich Prof. Dr. Awudu Abdulai, Prof. Dr. Uwe Latacz-Lohmann, Dr. Jerzy Michalek, Dr. Volker Saggau, Ursula Schwarz, Géraldine Jahnke, Carsten Struve, Anja Fleischhauer, Eva Krampe, Maren Harms, Mark Gregory, Rosa Gutschke sowie mehreren anonymen Gutachtern.

Meiner Familie und meinen Freunden danke ich von ganzem Herzen für ihre Unterstützung und das fortwährende Interesse an meiner Arbeit.

Kiel, im Juli 2006 — Arne Henningsen

Inhaltsverzeichnis

Kapitel 1

Einleitung und Zusammenfassung

In den meisten Ländern der Welt können starke staatliche Eingriffe im Agrarsektor beobachtet werden. Die Analyse der Auswirkungen dieser vielfältigen Agrarpolitiken ist sehr komplex und erfordert detaillierte Kenntnisse über die Angebots- und Nachfrageseite im Agrarsektor. Somit ist eine angemessene Modellierung des Produzenten- und Konsumentenverhaltens eine essentielle Voraussetzung für eine adäquate Wirkungsanalyse von Agrarpolitiken.

Die von der agrarökonomischen Forschung verwendeten Theorien und Methoden werden dabei häufig aus anderen Forschungsgebieten (z.B. der Mikroökonomie oder der Ökonometrie) übernommen und gegebenenfalls an die speziellen Rahmenbedingungen im Agrarsektor angepasst, teilweise auch selbst entwickelt. So ist beispielsweise das Unternehmens-Haushalts-Modell eine speziell in der Agrarökonomie entwickelte Theorie, die im Gegensatz zur klassischen mikroökonomischen Theorie die Interdependenzen zwischen Produktions- und Konsumentscheidungen in (landwirtschaftlichen) Familienbetrieben berücksichtigt.

Einen Schwerpunkt der agrarökonomischen Forschung stellen empirische Anwendungen dar. Da die Realität in der Regel sehr komplex ist, setzt eine adäquate Modellierung die Verwendung angemessener ökonomischer Theorien sowie fortgeschrittener empirischer Methoden voraus. Die daraus resultierenden Problemstellungen erweisen sich daher als nichttrivial und erfordern fundierte Kenntnisse sowohl in ökonomischer als auch in empirischer Methodik.

Die neun Aufsätze der hier vorliegenden Dissertation lassen sich in die oben beschriebenen Themengebiete einordnen. Eine Übersicht über die Forschungsbeiträge der einzelnen Aufsätze ist in Tabelle 1.1 dargestellt.

Die ersten vier Beiträge sind primär empirische Anwendungen, in denen das Konsumenten- und Produzentenverhalten mit adäquaten ökonometrischen Methoden und Programmierungsmodellen abgebildet wird. Der sechste und siebte Aufsatz bilden den Kern dieser Arbeit. Sie leisten Forschungsbeiträge sowohl im Bereich der ökonometrischen Methoden als auch in der mikroökonomischen Theorie und greifen diese anschließend in empirischen Anwendungen auf. Im fünften, achten und neunten Aufsatz werden schließlich spezielle ökonometrische Methoden entwickelt bzw. detailliert durchleuchtet.

Tabelle 1.1: Einordnung der Aufsätze

Aufsatz Nr.	empirische Anwendung Nachfrage	empirische Anwendung Produktion	Theorie	Methode
1	✓			
2	✓			
3		✓		
4		✓		
5		✓		✓
6	✓	✓	✓	✓
7	✓	✓	✓	✓
8				✓
9				✓

An Empirical Investigation of the Demand for Bananas in Germany

Die EU Bananenmarktordnung übt einen starken Einfluss auf die Bananenpreise in Deutschland aus. Auch von der für das Jahr 2006 geplanten Reform dieser Marktordnung sind deutliche Effekte zu erwarten. Eine detaillierte Analyse der Auswirkungen dieser staatlichen Eingriffe auf die Konsumenten und auf die weltweiten Handelsströme setzt u.a. genaue Kenntnisse über das Nachfrageverhalten der Konsumenten voraus. Vor diesem Hintergrund wird im ersten Aufsatz dieser Arbeit die Nachfrage der Konsumenten nach Bananen und anderen Obstsorten in Deutschland detailliert untersucht.

Ein weiterer Aspekt ist die eingehende Untersuchung des Gewohnheitsverhaltens der Konsumenten — d.h. möglicherweise verzögerte Anpassungsreaktionen — beim Bananenkonsum mit dynamischen Modellen. Zum einen ist das Wissen über dieses Verhalten beispielsweise für Bananenimporteure und Obsthändler relevant. Zum anderen wurde bei der Analyse mit statischen Modellen signifikante Autokorrelation festgestellt, sodass dynamische Modelle effizientere Schätzergebnisse liefern.

Insgesamt trägt dieser Aufsatz zu wesentlich detaillierteren Kenntnissen über die Bananennachfrage in Deutschland bei. Im Gegensatz zu den meisten bisherigen Untersuchungen auf diesem Gebiet (z.B. Wegner, 1989; Deodhar und Sheldon, 1995; Weiss, 1995; Herrmann, 1996; Herrmann und Sexton, 1999) werden (a) aktuelle Daten verwendet, (b) mikroökonomisch fundierte Nachfragemodelle eingesetzt, (c) dynamische Anpassungsreaktionen aufgrund von Gewohnheitsverhalten berücksichtigt, (d) verschiedene Haushaltstypen mit unterschiedlichem Einkommen untersucht und (e) Substitutionsbeziehungen zu anderem Frischobstarten detailliert analysiert.

Zur ökonometrischen Schätzung werden sowohl dynamische log-lineare Nachfragefunktionen als auch eine allgemeine dynamische Form des „Almost Ideal Demand Systems" (Deaton und Muellbauer, 1980a,b) verwendet. Die ökonometrische Schätzung dieses Modells erfolgt als Fehlerkorrekturmodell (Anderson und Blundell, 1982), bei dem die kurzfristigen Anpassungsreaktionen sowohl einem autoregressiven als auch einem „moving average"-Prozess folgen können. Durch die Annahme des „Multi-Stage-Budgetings" kann mit einer genesteten Struktur eine disaggregierte Analyse des Frischobstkonsums selbst bei moderater Datenverfügbarkeit durchgeführt werden.

Zusätzlich ermitteln wir die Determinanten der gesamten Frischobstnachfrage auf der oberen Stufe des genesteten Systems, um mit der Methode von Fan et al. (1995) die weitaus realistischeren „unbedingten" Nachfrageelastizitäten berechnen zu können. Weiterhin bestimmen wir — im Gegensatz zu vielen anderen Nachfrageanalysen — die Varianz-Kovarianz-Matrix der geschätzten Nachfrageelastizitäten, sodass die statistische Signifikanz bzw. die Genauigkeit dieser Elastizitäten festgestellt werden kann.

Durch die Verwendung von dynamischen Nachfragemodellen kann gezeigt werden, dass die Bananennachfrage signifikant durch das Gewohnheitsverhalten der Konsumenten beeinflusst wird. Die kurzfristigen Nachfrageanpassungen sind zumeist deutlich geringer als die langfristigen Reaktionen. Dabei passen ärmere Haushalte ihre Nachfrage wesentlich schneller den neuen Rahmenbedingungen an als reichere Haushalte. Während die Anpassungsreaktion ärmerer Haushalte schon nach zwei Monaten zu 96% erfolgt ist, dauert diese bei durchschnittlichen Haushalten drei Monate und bei reicheren Haushalten vier Monate. Neben der Geschwindigkeit unterscheidet sich auch die Höhe der Anpassung zwischen den Haushalten. So führt z.B. ein 1%iger Preisanstieg langfristig bei ärmeren Haushalten zu einem 0,6%igen Konsumrückgang und bei durchschnittlichen und reicheren Haushalten zu einem 0,4%igen Konsumrückgang. Weiter wurde nachgewiesen, dass andere Kategorien von Obst sowohl Brutto- als auch Nettosubstitute für Bananen sind.

Auf Grundlage dieser Ergebnisse ist es nun möglich, die Auswirkungen von Bananenmarktpolitiken auf deutsche Konsumenten und die weltweiten Handelsströme wesentlich genauer zu analysieren.

Die EU Bananenmarktordnung und die Nachfrage in Deutschland

Die Auswirkungen der 1993 eingeführten europäischen Bananenmarktordnung auf deutsche Verbraucher werden im zweiten Aufsatz analysiert. Während viele Studien (z.B. Kersten, 1994, 1995; Herrmann, 1996; Herrmann und Sexton, 1999; Kersten, 2000) die Effekte der Bananenmarktordnung auf den Welthandel ausführlich analysieren, werden die Auswirkungen auf die Konsumenten nur sehr oberflächlich untersucht. In dem hier vorliegenden Aufsatz wird nun diese Lücke für die deutschen Konsumenten geschlossen. Eine Besonderheit dieser Studie ist, dass drei verschiedene Haushaltstypen mit unterschiedlichem Einkommensniveau untersucht werden. Weiterhin werden die Effekte auf den Konsum von anderem Frischobst als mögliche Substitute detailliert analysiert. Als Grundlage dieser Untersuchung dienen die Ergebnisse aus dem ersten Aufsatz.

Durch die Einführung der EU Bananenmarktordnung stieg der Verbraucherpreis für Bananen in Deutschland um ca. 38% an. Damit hatten die deutschen Bananenkonsumenten im Vergleich zu anderen europäischen Konsumenten den höchsten Anstieg der Bananenpreise hinzunehmen. Unsere Untersuchung zeigt, dass dadurch der Bananenkonsum in den untersuchten Haushaltstypen um 12,5% bis 17,6% gesunken ist.[1] Aufgrund von Substitutionsbeziehungen steigt hingegen der Konsum von Kernobst und anderem Frischobst an. Die Ausgaben der Haushalte für Frischobst insgesamt erhöhen sich um 4,7% bis 6,1%. Allerdings entsprechen diese Ausgabenerhöhungen nur 0,05% (ärmere Haushalte) bis 0,036% (reichere Haushalte) der Gesamtausgaben der Haushalte.

Die EU Bananenmarktordnung hat somit zwar deutliche Auswirkungen auf die Bananennachfrage in Deutschland, aber die gesamten Frischobstausgaben der Haushalte steigen — zumindest relativ gesehen — nur minimal an.

Methodisches Vorgehen zur Modellierung der MTR-Beschlüsse

Der dritte Beitrag ist Teil einer Studie, die im Auftrag der Raiffeisen Hauptgenossenschaft Nord AG durchgeführt wurde (Henning et al., 2004). Das Ziel dieser Studie war es, die Auswirkungen des Mid-Term-Review (MTR) — der im Juni 2003 beschlossenen Reform der europäischen Agrarpolitik — auf den Agrarsektor und das Agribusiness in Schleswig-Holstein und Mecklenburg-Vorpommern quantitativ und möglichst disaggregiert zu untersuchen. In dem hier präsentierten Kapitel der Studie wird das methodische Vorgehen dargelegt, wobei der Schwerpunkt auf der Beschreibung des von mir entwickelten Gruppenhofmodells liegt.

Eine Untersuchung mit ökonometrischen Prognosemodellen setzt voraus, dass der Einfluss der betrachteten Maßnahmen in der Vergangenheit untersucht werden kann. Die zu untersuchende Agrarreform enthält jedoch Maßnahmen, die in der Vergangenheit noch nicht zu beobachten waren (z.B. Entkopplung, Cross-Compliance, starke Milchpreissenkung), sodass eine Analyse der Auswirkungen mit ökonometrischen Methoden in diesem Fall nicht ohne stark vereinfachende Annahmen durchgeführt werden konnte. Daher wurden die Aus-

[1] Bei größeren Preisänderungen entspricht die Mengenänderung nicht dem Produkt aus Preisänderung und Preiselastizität, da die Nachfragefunktionen des AIDS nicht-linear in Preisen sind.

wirkungen dieser Reform mit Programmierungsmodellen analysiert. Zum einen wurde ein regionalisiertes Gruppenhofmodell für Schleswig-Holstein und Mecklenburg-Vorpommern erstellt, zum anderen wurde das Modell CAPRI (Common Agricultural Policy Regional Impact Analysis) genutzt, welches vom Lehrstuhl für Agrarpolitik der Universität Bonn entwickelt wurde.

Mit Hilfe des regionalisierten Gruppenhofmodells werden die Auswirkungen des MTR auf Schleswig-Holstein und Mecklenburg-Vorpommern untersucht. Das für Schleswig-Holstein verwendete Gruppenhofmodell basiert auf 416 „typischen Betrieben“ (Gruppenhöfen), die sich in der Betriebsgröße, Betriebsausrichtung und regionalen Lage unterscheiden. In Mecklenburg-Vorpommern werden insgesamt 156 verschiedene Betriebe modelliert. Somit können disaggregierte Ergebnisse für verschiedene Betriebstypen und auch für verschiedene Regionen ermittelt werden. Für jeden Gruppenhof wird mit linearer Programmierung (LP) die optimale Produktionsstruktur unter verschiedenen Szenarien ermittelt. Um diverse produktionstechnische Gegebenheiten und agrarpolitische Bestimmungen zu berücksichtigen, sind die LP-Modelle stark disaggregiert und stellen jedem modellierten Betrieb ca. 1100 verschiedene Aktivitäten zur Auswahl, die in die Bereiche Ackerbau, Grünland, Tierproduktion und Prämien gegliedert werden können. Neben einem Base-Run-Szenario werden elf weitere Szenarien gerechnet, so dass die Auswirkungen von unterschiedlichen agrarpolitischen und ökonomischen Rahmenbedingungen analysiert werden können.

Zur Analyse der deutschland-, europa- und weltweiten Effekte wird das CAPRI-Modell verwendet. Es ist ein komparativ-statisches Simulationsmodell für den Agrarsektor, das aus einem Angebots- und einem Marktmodul besteht. In dem Angebotsmodul wird für jede der über 200 NUTS-II-Regionen in der EU-15 ein aggregiertes Programmierungsmodell gerechnet, in dem der Gewinn maximiert wird. Anschließend berechnet das Marktmodul markträumende Preise, die wiederum an das Angebotsmodul zurückgegeben werden. Dieser iterative Prozess wird so lange wiederholt, bis ein Gleichgewicht erreicht ist. Neben den Mitgliedsländer der EU werden zwölf weitere Regionen der Welt abgebildet. Auch mit dem CAPRI-Modell werden unterschiedliche Szenarien gerechnet.

Das im Rahmen dieser Studie entwickelte Gruppenhofmodell leistet im wesentlichen zwei Forschungsbeiträge. Zum einen wurde erstmals ein so detailliertes und disaggregiertes Agrarsektormodell für Schleswig-Holstein und Mecklenburg-Vorpommern entwickelt, und zum anderen wurden Methoden vorgestellt, mit denen die neuen Elemente der Agrarpolitik realitätsnah abgebildet werden können. Im Verlauf dieser Studie hat sich gezeigt, dass sich das Gruppenhof- und das CAPRI-Modell sehr gut ergänzen. Mit Hilfe des Gruppenhofmodells konnten disaggregierte Ergebnisse für bestimmte Betriebstypen und Gebiete in Schleswig-

Holstein und Mecklenburg-Vorpommern vorhergesagt werden. Das CAPRI-Modell eignete sich hingegen sehr gut, um die deutschland-, europa- und weltweiten Effekte zu prognostizieren (siehe Henning et al., 2004). Somit konnten die Auswirkungen des MTR auf allen für den Auftraggeber der Studie relevanten regionalen Ebenen vorhergesagt werden.

Economic Impact of the Mid-Term Review

Die Auswirkungen des Mid-Term-Review (MTR) auf Schleswig-Holstein werden im vierten Aufsatz untersucht. Dieser Beitrag basiert auf derselben Studie, aus der das vorangegangene Kapitel entnommen wurde. Allerdings beschränkt sich dieser Beitrag auf Schleswig-Holstein und geht nicht auf Mecklenburg-Vorpommern und das CAPRI-Modell ein. Zusätzlich werden in diesem Beitrag die Auswirkungen des MTR auf den betrieblichen Strukturwandel in Schleswig-Holstein untersucht.

Das verwendete Gruppenhofmodell wurde schon im vorangegangenem Beitrag beschrieben und wird daher hier nicht näher erläutert. Obwohl es sicherlich ein geeignetes Instrument ist, um die kurz- und mittelfristigen Auswirkungen des MTR auf die Produktionsentscheidungen und das Einkommen der Betriebe zu modellieren, ist es nicht in der Lage, die langfristigen Effekte auf den betrieblichen Strukturwandel zu untersuchen. Daher wurde ein weiteres Modell entwickelt, mit dem das Überleben von Betrieben und somit der Strukturwandel analysiert werden kann. Da das Ausscheiden von Betrieben hauptsächlich eine Frage der Hofnachfolge ist, verwenden wir eine bestehende Analyse der Hofnachfolgeentscheidungen in Schleswig-Holstein (Tietje, 2004). Mit den Ergebnissen dieser Studie können wir die Wahrscheinlichkeit der Hofnachfolge für verschiedene Betriebsgrößen und Betriebsausrichtungen sowie bei unterschiedlichen landwirtschaftlichen Einkommen abschätzen. Die Gewinne der Betriebe unter verschiedenen ökonomischen und agrarpolitischen Rahmenbedingungen ermitteln wir mit dem oben beschriebenen Gruppenhofmodell und setzen diese in das Strukturwandelmodell ein, sodass wir auch die Wahrscheinlichkeiten der Hofnachfolge unter verschiedenen Rahmenbedingungen ermitteln können.

Die Ergebnisse des Gruppenhofmodells zeigen, dass es auf aggregierter Ebene zu keiner dramatischen Änderung der Produktionsstruktur kommen wird. Während auf Standorten mit guten Bodenqualitäten die Bullenmast verringert und der Marktfruchtbau ausgedehnt wird, wird auf Standorten mit schlechter Bodenqualität der Futterbau extensiviert und der Marktfruchtbau eingeschränkt. Aufgrund der Milchpreissenkungen verringert sich das Einkommen der Milchviehbetriebe — im Gegensatz zu den meisten anderen Betriebstypen — deutlich (um bis zu 37%). Obwohl der MTR zum Teil einen deutlichen Einfluss auf die landwirtschaftlichen Einkommen hat, zeigen die Ergebnisse des Strukturwandelmodells, dass die

Geschwindigkeit des Strukturwandels nur geringfügig zunimmt. Unabhängig vom MTR wird die Anzahl der kleineren Betriebe deutlich abnehmen und die durchschnittliche Betriebsgröße zunehmen.

Unser Gruppenhofmodell hat im Einklang mit bestehenden aggregierten Modellen wie beispielsweise CAPRI prognostiziert, dass die Auswirkungen des MTR auf Schleswig-Holstein und Mecklenburg-Vorpommern insgesamt eher gering ausfallen. Allerdings konnte mit unserem disaggregierten Gruppenhofmodell zusätzlich gezeigt werden, dass trotz geringer Auswirkungen auf aggregierter Ebene deutliche Auswirkungen für bestimmte Betriebstypen und für bestimmte Regionen zu erwarten sind.

Analyse von Transaktionskosten auf dem ländlichen Kreditmarkt in Polen

Die Funktionsfähigkeit der ländlichen Kreditmärkte ist eine essentielle Voraussetzung für die Wettbewerbsfähigkeit der landwirtschaftlichen Produktion. Da der landwirtschaftliche Sektor in vielen Entwicklungs- und Transformationsländern noch eine verhältnismäßig bedeutsame gesamtwirtschaftliche Stellung hat, hat auch die Funktionsfähigkeit der ländlichen Kreditmärkte in diesen Ländern eine große gesamtwirtschaftliche Bedeutung. Allerdings ist beispielsweise in der polnischen Landwirtschaft nur eine sehr geringe Investitionstätigkeit zu beobachten. Daher versucht der polnische Staat, durch umfangreiche Subventionsprogramme die Investitionstätigkeit der landwirtschaftlichen Betriebe zu erhöhen. In der Literatur wird der Zustand des ländlichen Kreditmarktes in Polen kontrovers diskutiert. Während die geringe Investitionstätigkeit von einigen mit einem Mangel an profitablen Investitionsprojekten begründet wird (Petrick, 2000), führen andere sie auf einen erschwerten Zugang der Landwirte zum ländlichen Kreditmarkt zurück (Moosburger et al., 1999; Puslecki, 2000).

Daher entwickeln wir eine Methode zur Quantifizierung von Transaktionskosten, die aufgrund institutioneller Mängel auf ländlichen Kreditmärkten vorliegen können (vgl. Koester, 2001). Zunächst wird eine quadratische Produktionsfunktion ökonometrisch geschätzt, um das Grenzprodukt des eingesetzten Kapitals zu ermitteln. Die Konkavität der Produktionsfunktion wurde dabei mit einer zweistufigen Minimum-Distance Schätzung (Koebel et al., 2003) erzwungen, damit die ermittelten Grenzprodukte aus einer mikroökonomisch konsistenten Technologie abgeleitet werden können.[2] Mit Hilfe des Annuitätenfaktors kann dann daraus die Verzinsung des eingesetzten Kapitals berechnet werden. Eine positive Differenz zwischen der ermittelten Kapitalverzinsung und dem tatsächlich gezahlten Zinssatz

[2] Zur Konkavität bzw. Quasikonkavität der Produktionsfunktion siehe Diskussion zu diesem Aufsatz in der Schlussbetrachtung (Seite 284).

kann dabei sowohl durch Transaktionskosten auf den Kreditmärkten als auch durch Risikoaversion zustande kommen. Durch Simulation von Risikoaufschlägen können wir dennoch Aussagen über mögliche Transaktionskosten ableiten. Diese Quantifizierung der Transaktionskosten erfolgt somit analog zu der komparativen Vorgehensweise von Williamson (1985, 2000).

Diese Methode wurde schließlich auf Daten polnischer Betriebe angewendet. Um die Ergebnisse dieser Analyse besser einordnen zu können, wurde das gleiche Verfahren mit Daten von schleswig-holsteinischen Betrieben wiederholt. Unsere Simulationsergebnisse zeigen, dass Transaktionskosten selbst dann noch einen bedeutenden Einfluss auf die ländlichen Kreditmärkte und die Kapitalausstattung der landwirtschaftlichen Betriebe in Polen haben, wenn wir den polnischen Landwirten eine unrealistisch hohe Risikoaversion unterstellen. Für die schleswig-holsteinischen Betriebe konnten wir hingegen zeigen, dass bei einem moderaten Risikoaufschlag keine Transaktionskosten mehr vorliegen.

Mit der in diesem Aufsatz entwickelten Methode konnte somit gezeigt werden, dass die Investitionstätigkeit in der polnischen Landwirtschaft durch Transaktionskosten auf dem ländlichen Kreditmarkt gehemmt wird.

Modeling Farm Households' Price Responses

In der landwirtschaftlichen Entwicklungsökonomie wurde schon vor langem erkannt, dass ländliche Märkte häufig unvollkommen oder gar nicht vorhanden sind (z.B. de Janvry et al., 1991). Die Produktions- und Konsumentscheidungen von Familienbetrieben, die aufgrund dieser Unvollkommenheiten nicht an Märkten teilnehmen, orientieren sich daher nicht an exogenen Marktpreisen, sondern an internen Schattenpreisen. Dadurch entstehen Interdependenzen zwischen Produktions- und Konsumentscheidungen, sodass herkömmliche mikroökonomische Modelle nicht mehr anwendbar sind. Stattdessen wurden Unternehmens-Haushalts-Modelle (UHMs) entwickelt, die die Zusammenhänge zwischen Produktions- und Konsumentscheidungen explizit berücksichtigen (de Janvry et al., 1991). Während UHMs zunächst verwendet wurden, um paradox erscheinende Anpassungsreaktionen von landwirtschaftlichen Familienbetrieben zu erklären (z.B. Lopez, 1984; Strauss, 1986; de Janvry et al., 1991, 1992), analysieren neuere UHM-Studien zumeist die Rolle von Transaktionskosten und Institutionen (z.B. Goetz, 1992; Key et al., 2000; Carter und Yao, 2002; Carter und Olinto, 2003; Vakis et al., 2003; Vance und Geoghegan, 2004). Dabei werden im Allgemeinen nur fixe und proportionale Transaktionskosten berücksichtigt.

Ausgehend von diesen vorhandenen UHM-Ansätzen habe ich in den letzten Jahren diverse Weiterentwicklungen vorgenommen, die in diesem und auch im folgenden Aufsatz vorgestellt werden.[3] Zunächst werden die beiden oben aufgeführten Forschungsrichtungen verbunden, indem die Anpassungsreaktionen von landwirtschaftlichen Familienbetrieben unter Berücksichtigung von Transaktionskosten auf Arbeitsmärkten untersucht werden. Dabei analysieren wir nicht nur fixe und proportionale, sondern auch nicht-proportionale variable Transaktionskosten, sodass unser Modell bei vielen verschiedenen Arten von Marktunvollkommenheiten anwendbar ist. Zusätzlich berücksichtigen wir Heterogenität der Arbeitskräfte, denn in der Regel gibt es gerade auf Arbeitsmärkten große Qualitätsunterschiede. Diese Erweiterungen des klassischen UHMs erlauben uns, Anpassungsreaktionen von Familienbetrieben unter verschiedenen Arbeitsmarktregimen zu untersuchen. Dies beinhaltet auch den Fall, in dem Arbeit gleichzeitig angeboten und nachgefragt wird, der in den meisten bisherigen Studien ausgeschlossen ist.

Ein weiterer Schwerpunkt dieses Beitrags liegt auf der konsistenten ökonometrischen Schätzung des UHMs. Im Gegensatz zu den meisten bisherigen Studien, in denen nicht-separable UHMs auf „reduced form“-Gleichungen basieren, verwenden wir ein vollständiges UHM mit theoretisch konsistenten flexiblen Funktionsformen auf Produktions- und Konsumseite. Durch den Einsatz flexibler Funktionsformen stellen wir sicher, dass wir die ökonometrischen Schätzungen nicht durch unerwünschte Restriktionen beschränken. Darüber hinaus gewährleistet die Verwendung theoretisch konsistenter Funktionen, dass das modellierte Anpassungsverhalten nicht der mikroökonomischen Theorie widerspricht.

So wird z.B. die Produktionstechnologie der landwirtschaftlichen Betriebe mit einer „Symmetric Normalized Quadratic“-Profitfunktion modelliert, wobei die von der mikroökonomischen Theorie geforderte Konvexität dieser Funktion mit einer aktuellen ökonometrischen Methode von Koebel et al. (2003) erreicht wird. Die Konsumentscheidungen der Haushalte werden mit einem „Almost Ideal Demand System“ (Deaton und Muellbauer, 1980a,b) abgebildet. Allerdings haben wir es im Gegensatz zu den meisten bestehenden empirischen Anwendungen bei der ökonometrischen Schätzung nicht linear approximiert, sondern als vollständiges nicht-lineares Modell mit einem iterativen linearen Verfahren von Blundell und Robin (1999) konsistent ökonometrisch geschätzt. Wie im achten Aufsatz gezeigt wurde,

[3] Diese beiden Aufsätze gehören sicherlich zu den wichtigsten Beiträgen dieser Arbeit, denn diesen Weiterentwicklungen habe ich den größten Teil meiner Dissertationszeit gewidmet. Einen ersten Vorläufer dieser beiden Aufsätze habe ich im Juli 2002 auf der Konferenz der „American Agricultural Economics Association“ in Long Beach, Kalifornien, vorgestellt. Auf der Konferenz der „International Association of Agricultural Economists“ im August 2003 in Durban, Südafrika, konnte ich eine verbesserte Version präsentieren. Eine weitere überarbeitete Fassung dieser Aufsätze habe ich im April 2005 auf einem Seminar der „European Association of Agricultural Economists“ in Wye, England, vorgetragen. Im August 2005 haben wir dann eine nochmals verbesserte Version dieses Artikels beim „American Journal of Agricultural Economics“ eingereicht. Nach Berücksichtigung der Hinweise der Gutachter haben wir im Februar 2006 eine überarbeitete Fassung für die „zweite Runde“ vorgelegt. Nach einer weiteren Überarbeitung haben wir schließlich im Mai 2006 diese hier abgedruckte vorläufige Endfassung eingereicht.

führt dieses iterative Verfahren zu weitaus genaueren Schätzergebnissen als eine lineare Approximation des Modells. Bei der ökonometrischen Schätzung sowohl der Profitfunktion als auch des Nachfragesystems musste beachtet werden, dass infolge von Arbeitsmarktunvollkommenheiten der Preis der Arbeit bzw. der Freizeit nicht exogen vorgegeben ist, sondern sich endogen ergibt. Um eine Verzerrung der Regressionsergebnisse zu verhindern, haben wir diese Gleichungssysteme mit der Instrument-Variablen-Methode „Three-Stage Least Squares“ (3SLS) ökonometrisch geschätzt.

Bei der ökonometrischen Analyse der Arbeitsmärkte mussten drei mögliche Probleme umgangen werden. Zum einem ist davon auszugehen, dass die Familien ihre Arbeitsangebots- und Arbeitsnachfrageentscheidungen nicht getrennt, sondern simultan treffen. Dieses Problem haben wir durch die simultane ökonometrische Schätzung der Arbeitsangebots- und Arbeitsnachfrageentscheidungen mit einem bivariaten probit-Modell gelöst. Zum zweiten können die Lohngleichungen für angebotene und nachgefragte Arbeit nur jeweils für diejenigen Familienbetriebe geschätzt werden, die auch tatsächlich Arbeit anbieten bzw. nachfragen. Dadurch kann eine so genannte „sample selection“-Verzerrung auftreten, die im Allgemeinen mit einem zweistufigen Verfahren von Heckman (1976) konsistent ökonometrisch geschätzt werden kann. Allerdings ist dieses Schätzverfahren nur für univariate probit-Modelle ausgelegt, sodass wir eine Erweiterung dieses Modells für bivariate probit-Modelle vornehmen mussten. Schließlich muss bei den ökonometrischen Schätzungen der Lohngleichungen berücksichtigt werden, dass einige Regressoren endogene Entscheidungsvariablen sind, und deshalb die geschätzten Koeffizienten verzerrt sein können. Dieses Problem konnten wir mit einem Schätzverfahren von Lee et al. (1980) für simultane Gleichungsmodelle bei Selektivität umgehen. Allerdings mussten wir auch dieses Verfahren für bivariate probit-Modelle erweitern. Die Ergebnisse dieses Schätzverfahrens erlauben schließlich einen einfachen statistischen Test, mit dem die Separabilität zwischen Produktions- und Konsumentscheidungen empirisch überprüft werden kann.

Unsere theoretische Untersuchung zeigt, dass in unserem UHM-Ansatz Nicht-Separabilität nicht nur bei autarken Haushalten auftritt, sondern auch, wenn Haushalte an (unvollkommenen) Arbeitsmärkten teilnehmen. Eine komparativ statische Analyse zeigt, dass im Falle von unvollkommenen Arbeitsmärkten die Richtung der meisten Preisanpassungsreaktionen theoretisch nicht bestimmt ist, da gegenläufige Schattenpreiseffekte auftreten. Weiterhin unterscheiden sich die Anpassungsreaktionen für separable und nicht-separable UHMs. Dies zeigt, dass Marktunvollkommenheiten einen Einfluss auf die Anpassungsreaktionen von Familienbetrieben haben.

Unsere ökonometrische Analyse mit detaillierten Haushalts- und Betriebsdaten von polnischen Familienbetrieben zeigt, dass signifikante Unvollkommenheiten auf den ländlichen Arbeitsmärkten in Polen vorliegen. Das Ausmaß der Unvollkommenheiten auf den Arbeitsmärkten hat dabei einen signifikanten Einfluss auf die Preisanpassungsreaktionen der Familienbetriebe. Die Ergebnisse unseres Ansatzes unterscheiden sich von denen herkömmlicher separabler und nicht-separabler UHMs signifikant.

Unser UHM-Ansatz stellt eine Erweiterung und Verallgemeinerung bisheriger Ansätze dar. Weiterhin stellen wir ökonometrische Methoden vor, mit denen nicht-separable Unternehmens-Haushalts-Modelle konsistent ökonometrisch geschätzt werden können. Die empirische Anwendung dieses Ansatzes zeigt, dass die Berücksichtigung von nicht-proportionalen variablen Transaktionskosten und Heterogenität einen entscheidenden Einfluss auf die Ergebnisse hat. Daher ist davon auszugehen, dass es in vielen Fällen die Realität besser abbilden kann als herkömmliche unflexiblere UHMs.

Taxation of the Farm Household and Imperfect Labor Markets

Ein zentrales Ergebnis der klassischen finanzwissenschaftlichen Literatur ist, dass eine optimale Besteuerung aus einer Kombination aus Konsum- und Einkommensteuern besteht, da von diesen Standard-Steuern keine direkten Produktionseffekte ausgehen (Ramsey, 1927). Allerdings weicht in vielen Ländern die Besteuerung landwirtschaftlicher Familienbetriebe von der regulären Besteuerung ab, und es werden stattdessen spezielle landwirtschaftliche Steuern erhoben. Da von den meisten landwirtschaftlichen Steuern im Gegensatz zu Standard-Steuern direkte Produktionseffekte ausgehen, wird davon ausgegangen, dass erstere volkswirtschaftlich ineffizient sind.

Allerdings kommen neuere Studien zur Besteuerung der Landwirtschaft, die Unvollkommenheiten auf ländlichen Märkten berücksichtigen, teilweise zu anderen Ergebnissen als die klassische finanzwissenschaftliche Literatur (Hoff und Stiglitz, 1993; Hoff, 1993). Zum einen gelten die allgemeinen Effizienztheoreme der Besteuerung nur bei vollkommenen Märkten, sodass diese bei Marktunvollkommenheiten nicht mehr angewendet werden können. Zum anderen beeinflussen Marktunvollkommenheiten die Reaktionen auf Steuerpolitiken, sodass auch Standard-Steuern direkte Produktionseffekte haben können und somit zu Ineffizienzen führen können. Allerdings sind diese neueren Studien ausschließlich theoretisch und beschränken sich entweder nur auf landwirtschaftliche Steuern (z.B. Hoff, 1993) oder nur auf Standard-Steuern (z.B. Chambers und Lopez, 1987).

Dagegen wird in diesem Beitrag gleichzeitig die Wirkung von speziellen landwirtschaftlichen Steuern und Standard-Steuern auf landwirtschaftliche Familienbetriebe sowohl theoretisch als auch empirisch untersucht. Zur theoretischen Analyse verwenden wir das im vorherigen Beitrag entwickelte Unternehmens-Haushalts-Modell (UHM) und leiten die komparative Statik für diverse landwirtschaftliche Steuern und Standard-Steuern ab. Die mit unserem Modell ermittelten Effekte der Steuern bestätigen die Ergebnisse der neueren Studien und zeigen, dass von allen untersuchten Steuerarten Produktionseffekte ausgehen. Daher sind Standard-Steuern theoretisch nicht mehr den landwirtschaftlichen Steuern überlegen.

Neben der theoretischen komparativ-statischen Analyse wird auch eine empirische Untersuchung vorgenommen, um der Frage nachzugehen, in wie weit und unter welchen Umständen theoretisch nachgewiesene Auswirkungen von Marktunvollkommenheiten überhaupt bedeutende empirische Effekte aufweisen. Diese Frage ist insbesondere für die praktische Gestaltung von landwirtschaftlichen Steuersystemen relevant.

Auch unsere empirischen Untersuchungen basieren auf den im vorherigen Beitrag vorgestellten ökonometrischen Schätzungen. Aus den Schätzergebnissen berechnen wir mit Hilfe der komparativen Statik „Steuerelastizitäten“, die die prozentualen Anpassungsreaktionen bei einer Erhöhung der Steuern um einen Prozentpunkt angeben. Diese „Steuerelastizitäten“ geben wir für verschiedene Arbeitsmarktregime an, sodass der Einfluss des Arbeitsmarktregimes bzw. der Arbeitsmarktunvollkommenheiten auf die Anpassungsreaktionen deutlich wird.

Im Gegensatz zu unserer theoretischen Analyse bekräftigt unsere empirische Analyse polnischer Familienbetriebe zum Teil die klassische finanzwissenschaftliche Literatur. Bei unvollkommenen aber existenten Arbeitsmärkten haben Konsum- und Einkommensteuern vernachlässigbare Produktionseffekte, während Vermarktungs- oder Vorleistungssteuern bedeutende Auswirkungen auf die Produktionsstruktur aufweisen. Dagegen lösen bei fehlenden Arbeitsmärkten auch die Konsum- und Einkommensteuer bemerkenswerte Produktionseffekte aus. Die Landsteuer verursacht in beiden Szenarien beachtliche Produktionsanpassungen.

Die Ergebnisse zeigen, dass es bei unvollkommenen Märkten keine einfache Faustregel zur optimalen Besteuerung von landwirtschaftlichen Familienbetrieben gibt. Somit erfordert die optimale Gestaltung von landwirtschaftlichen Steuersystemen eine umfassende quantitative Wirkungsanalyse von verschiedenen Steuerpolitiken, die beispielsweise mit dem hier vorgestellten UHM vorgenommen werden kann.

How to Estimate the “Almost Ideal Demand System”

Im achten Aufsatz werden verschiedene Verfahren zur ökonometrischen Schätzung des „Almost Ideal Demand Systems“ (AIDS) analysiert und verglichen. Das AIDS gehört zu den am häufigsten in empirischen Nachfrageanalysen verwendeten Nachfragesystemen, da es fast alle wünschenswerten Eigenschaften eines Nachfragesystems aufweist. Um eine komplizierte nicht-lineare Schätzung des AIDS zu umgehen, verwenden die meisten empirischen Studien eine lineare Approximation des AIDS (LA-AIDS), in der der Translog-Preisindex durch den Stone-Preisindex ersetzt wird. Allerdings führt die Approximation durch den Stone-Preisindex zu einer sehr schlechten Approximation der Koeffizienten des originalen AIDS. Darüber hinaus treten bei der ökonometrischen Schätzung des LA-AIDS durch die Approximation mehrere zusätzliche Probleme auf. Hinzu kommt, dass das LA-AIDS kein theoretisch konsistentes Nachfragesystem mehr ist. Einige Wissenschaftler haben versucht, diese Probleme dadurch zu umgehen, dass sie spezielle Formeln zur Berechnung der Nachfrageelastizitäten ableiten oder andere Preisindizes verwenden (z.B. Green und Alston, 1990; Moschini, 1995). Da aber keine dieser Vorgehensweisen alle Schätzprobleme löst und das LA-AIDS ein theoretisch inkonsistentes Nachfragesystem bleibt, haben andere Forscher iterative lineare Verfahren zur ökonometrischen Schätzung des nicht-linearen AIDS vorgeschlagen (Browning und Meghir, 1991; Michalek und Keyzer, 1992; Blundell und Robin, 1999)[4].

In diesem Aufsatz werden nun diese verschiedenen Schätzverfahren mit einer Monte-Carlo-Simulation analysiert und verglichen. Im Gegensatz zu bestehenden Monte-Carlo-Analysen des AIDS und LA-AIDS (Buse, 1993, 1994; Buse und Chan, 2000) weist die hier vorgestellte Untersuchung mehrere Verbesserungen und Erweiterungen auf. (a) Die simulierten Störterme werden mit der in diesem Fall realistischeren Dirichlet-Verteilung erzeugt. (b) Die Simulation wird unter zusätzlichen Rahmenbedingungen (z.B. Endogenität der Gesamtausgaben) durchgeführt, die für einige empirische Nachfrageanalysen realistischer sind. (c) Zusätzlich zu den Koeffizienten und Nachfrageelastizitäten wird auch die Genauigkeit von Wohlfahrtseffekten untersucht. (d) Die störanfällige Schätzung der Konstante des Translog-Preisindexes (α_0) wird analysiert. (e) Iterative lineare Schätzverfahren werden systematisch berücksichtigt. (f) Die Effekte einer Mittelwertbereinigung der Preise und der Ausgaben werden betrachtet.

Die Ergebnisse dieser Untersuchung zeigen, dass mit der Schätzung eines einfachen LA-AIDS mit Tornqvist- oder Laspeyres-Preisindex akzeptable Schätzer für Nachfrageelastizitäten ermittelt werden können. Wenn Preise und Ausgaben vor der Schätzung mittelwertbereinigt werden, können die Ergebnisse sogar zur Wohlfahrtsanalyse verwendet werden. Allerdings führt eine iterative lineare Schätzung des theoretisch konsistenten nicht-linearen AIDS zu wesentlich genaueren Schätzergebnissen, sodass die Schätzung des theoretisch inkonsistenten LA-AIDS auch ohne nicht-lineare Schätzverfahren vermieden werden kann. Bei der Schätzung des nicht-linearen AIDS — egal ob iterativ linear oder nicht-linear — ist es wichtig, die Konstante des Translog-Preisindizes (α_0) nicht a priori festzulegen, sondern an-

[4]Siehe hierzu die Diskussion zu diesem Aufsatz in der Schlussbetrachtung (Seite 287).

hand der Schätzergebnisse auszuwählen. Im Gegensatz zu anderen Veröffentlichungen zeigen unsere Ergebnisse, dass eine Mittelwertbereinigung der Preise und Gesamtausgaben nicht zu verzerrten Schätzergebnissen führt. Daher kann dieses Verfahren angewendet werden, um die Berechnung von Nachfrageelastizitäten deutlich zu vereinfachen.

systemfit: Simultaneous Equation Systems in R

Viele statistische und ökonometrische Analysen basieren auf Modellen, die nicht nur aus einer einzelnen Gleichung, sondern aus einem ganzen Gleichungssystem bestehen. Im Bereich der Mikroökomie sind dies u.a. Nachfragesysteme auf Konsumentenseite (siehe z.B. den ersten, zweiten, sechsten, siebten und achten Aufsatz dieser Arbeit) sowie Kosten- oder Profitfunktionen auf Unternehmensseite (siehe z.B. den sechsten und siebten Aufsatz dieser Arbeit). Im Fall von Mehrgleichungsmodellen ist davon auszugehen, dass die Störterme der Gleichungen korreliert sind. Daher erfordert eine effiziente Schätzung dieser Modelle eine simultane Schätzung aller Gleichungen (z.B. als „Seemingly Unrelated Regression“). Wenn weiterhin einige erklärende Variablen in anderen Gleichungen als endogene Variable auftreten oder aus anderen Gründen mit den Störtermen korreliert sind, müssen zusätzlich Instrumentvariablen-Methoden verwendet werden, um unverzerrte Schätzer zu erhalten (z.B. „3-Stage Least Squares“).

In diesem Aufsatz wird das Softwarepaket „systemfit“ vorgestellt, das die ökonometrische Schätzung von Mehrgleichungsmodellen in der Statistik-Software „R“ (R Development Core Team, 2005) ermöglicht. Es stellt u.a. die Schätzmethoden „Ordinary Least Squares“, „Weighted Least Squares“, „Seemingly Unrelated Regression“, „2-Stage Least Squares“, „Weighted 2-Stage Least Squares“ und „3-Stage Least Squares“ bereit. Dabei können die zu schätzenden Parameter durch beliebige lineare Restriktionen beschränkt werden. Weiterhin stellt „systemfit“ Funktionen bereit, um Parameter-Restriktionen mit der F-, Wald- oder Likelihood-Ratio-Statistik zu testen, und um die Konsistenz des „3-Stage Least Squares“-Schätzers mit dem Hausman-Test zu überprüfen. Die Genauigkeit und Zuverlässigkeit von „systemfit“ wird festgestellt, indem die Ergebnisse von veröffentlichten Studien erfolgreich repliziert werden können.

Im Vergleich zu bestehender Software ist das Paket „systemfit“ wesentlich flexibler, denn die Benutzer können viele Details der Schätzung beeinflussen. Da sowohl „systemfit“ als auch „R“ Open-Source-Software sind, ist es im Gegensatz zu proprietärer Software für jeden möglich, die verwendeten Algorithmen zu überprüfen („peer review“) und gegebenenfalls zu erweitern. Dies ist gerade in der Wissenschaft ein großer Vorteil. Schließlich ist die schnelle und weite Verbreitung dieses Softwarepakets sicherlich ein Indiz für dessen Relevanz.[5]

Literaturverzeichnis

Anderson, G. J. und R. W. Blundell (1982): Estimation and Hypothesis Testing in Dynamic Singular Equation Systems. In: Econometrica, 50(6):1559–1571.

Blundell, R. W. und J. M. Robin (1999): Estimation in Large and Disaggregated Demand Systems: An Estimator for Conditionally Linear Systems. In: Journal of Applied Econometrics, 14:209–232.

Browning, M. und C. Meghir (1991): The Effects of Male and Female Labor Supply on Commodity Demands. In: Econometrica, 59:925–951.

Buse, A. (1993): Evaluating the Linearized Almost Ideal Demand System. Research Paper No. 93-18, Department of Economics, University of Alberta, Edmonton.

Buse, A. (1994): Evaluating the Linearized Almost Ideal Demand System. In: American Journal of Agricultural Economics, 76:781–793.

Buse, A. und W. H. Chan (2000): Invariance, Price Indices and Estimation in Almost Ideal Demand Systems. In: Empirical Economics, 25:519–539.

Carter, M. R. und P. Olinto (2003): Getting Institutions "Right" for Whom? Credit Constraints and the Impact of Property Rights on the Quantity and Composition of Investment. In: American Journal of Agricultural Economics, 85(1):173–186.

Carter, M. R. und Y. Yao (2002): Local versus Global Separability in Agricultural Household Models: The Factor Price Equalization Effect of Land Transfer Rights. In: American Journal of Agricultural Economics, 84(3):702–715.

Chambers, R. G. und R. E. Lopez (1987): Tax Policies and the Financially Constrained Farm Household. In: American Journal of Agricultural Economics, 69(2):369–377.

de Janvry, A., M. Fafchamps und E. Sadoulet (1991): Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained. In: The Economic Journal, 101(409):1400–1417.

[5]Eine Recherche unter http://www.google.com findet zur Zeit (4. 5. 2006) ca. 18.100 Internetseiten, auf denen die Begriffe „systemfit“ und „R“ gleichzeitig vorkommen.

de Janvry, A., E. Sadoulet, M. Fafchamps und M. Raki (1992): Structural Adjustment and the Peasantry in Morocco: A Computable Household Model. In: European Review of Agricultural Economics, 19(4):427–453.

Deaton, A. und J. Muellbauer (1980a): An Almost Ideal Demand System. In: The American Economic Review, 70:312–326.

Deaton, A. und J. Muellbauer (1980b): Economics and Consumer Behaviour. Cambridge University Press, Cambridge.

Deodhar, S. Y. und I. Sheldon (1995): Is Foreign Trade (Im)Perfectly Competitive?: An Analysis of the German Market for Banana Imports. In: Journal of Agricultural Economics, 46(3):336–348.

Fan, S., E. J. Wailes und G. L. Cramer (1995): Household Demand in Rural China: A Two-Stage LES-AIDS Model. In: American Journal of Agricultural Economics, 77:54–62.

Goetz, S. J. (1992): A Selectivity Model of Household Food Marketing Behavior in Sub-Sahara Africa. In: American Journal of Agricultural Economics, 74(2):444–452.

Green, R. und J. M. Alston (1990): Elasticities in AIDS models. In: American Journal of Agricultural Economics, 72:442–445.

Heckman, J. (1976): The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models. In: Annals of Economic and Social Measurement, 5:475–492.

Henning, C. H. C. A., A. Henningsen, C. Struve und J. Müller-Scheessel (2004): Auswirkungen der Mid-Term-Review-Beschlüsse auf den Agrarsektor und das Agribusiness in Schleswig-Holstein und Mecklenburg-Vorpommern. Agrarwirtschaft, Sonderheft 178, Agrimedia, Bergen/Dumme.

Herrmann, R. (1996): Economic Implications of the New European Banana Market Regime: The Case of Germany. Discussion Papers in Agricultural Economics 38, Department of Agricultural Economics, University of Gießen, Gießen.

Herrmann, R. und R. Sexton (1999): Redistributive Implications of a Tariff-rate Quota Policy: How Market Structure and Conduct Matter. Discussion Papers in Agricultural Economics 52, Department of Agricultural Economics, University of Gießen, Gießen.

Hoff, K. (1993): Land Taxes, Output Taxes, and Sharecropping: Was Henry George Right? In: Hoff, K., A. Braverman und J. E. Stiglitz (Hrsg.), The Economics of Rural Organization. Oxford University Press, New York, 332–351.

Hoff, K. und J. E. Stiglitz (1993): Consequences of Limited Risk Markets and Imperfect Information for the Design of Taxes and Transfers: An Overview. In: Hoff, K., A. Braverman und J. E. Stiglitz (Hrsg.), The Economics of Rural Organization. Oxford University Press, New York, 323–331.

Kersten, L. (1994): Wirkungen der EWG-Marktordnung für Bananen auf Handel und Preisbildung. In: Agrarwirtschaft, 43(8/9):319–327.

Kersten, L. (1995): Impacts of the EU Banana Market Regulation on International Competition, Trade and Welfare. In: European Review of Agricultural Economics, 22:321–335.

Kersten, L. (2000): Zur Revision der Bananenmarktordnung. Unveröffentlicht. Zusammenfassung veröffentlicht im Jahresbericht 1999 der FAL Braunschweig-Völkenrode, Kapitel 5.1, S. 136-137.

Key, N., E. Sadoulet und A. de Janvry (2000): Transaction Costs and Agricultural Household Supply Response. In: American Journal of Agricultural Economics, 82(2):245–259.

Koebel, B., M. Falk und F. Laisney (2003): Imposing and Testing Curvature Conditions on a Box-Cox Cost Function. In: Journal of Business and Economic Statistics, 21(2):319–335.

Koester, U. (2001): Agricultural Finance and Institutional Reforms in Transition Economies: The 1990s and Challenges Ahead. In: Quarterly Journal of International Agriculture, 40(4):301–323.

Lee, L., G. S. Maddala und R. P. Trost (1980): Asymptotic Covariance Matrices of Two-stage Probit and Two-stage Tobit Methods for Simultaneous Equations Models with Selectivity. In: Econometrica, 48(2):491–503.

Lopez, R. (1984): Estimating Labor Supply and Production Decisions of Self-Employed Farm Producers. In: European Economic Review, 24:61–82.

Michalek, J. und M. A. Keyzer (1992): Estimation of a Two-stage LES-AIDS Consumer Demand System for Eight EC Countries. In: European Review of Agricultural Economics, 19(2):137–163.

Moosburger, A., A. Balmann und M. Odening (1999): Struktur- und Effizienzwirkungen der einzelbetrieblichen Investitionsförderung in Polen. In: Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., 35:360–365.

Moschini, G. (1995): Units of Measurement and the Stone Index in Demand System Estimation. In: American Journal of Agricultural Economics, 77:63–68.

Petrick, M. (2000): Are Polish Farmers Credit Constrained? Theoretical Considerations and First Results of a Survey in Western Pomerania (Poland). Paper Presented at the International Scientific Conference 'Regional Structural Transformation of Agriculture and Rural Areas with Regard to Poland's Integration with the EU', Krakow, Poland, 7.–8. 12. 2000.

Puslecki, Z. (2000): Polens Landwirtschaft zwischen Beharrung und EU-Beitritt. In: WeltTrends - Zeitschrift für internationale Politik und vergleichende Studien, 27:95–107.

R Development Core Team (2005): R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. In: http://www.R-project.org. ISBN 3-900051-07-0.

Ramsey, F. (1927): A Contribution to the Theory of Taxation. In: Economic Journal, 37:47–61.

Strauss, J. (1986): The Theory and Comparative Statics of Agricultural Household Models: A General Approach. In: Singh, I., L. Squire und J. Strauss (Hrsg.), Agricultural Household Models: Extensions, Applications, and Policy. Johns Hopkins University Press, Baltimore, 71–91.

Tietje, H. (2004): Hofnachfolge in Schleswig-Holstein. Dissertation, Christian-Albrechts-Universität, Kiel, Germany.

Vakis, R., E. Sadoulet und A. de Janvry (2003): Measuring Transactions Costs from Observed Behavior: Market Choices in Peru. http://are.berkeley.edu/~sadoulet/papers/TC-wp.pdf.

Vance, C. und J. Geoghegan (2004): Modeling the Determinants of Semi-Subsistent and Commercial Land Uses in an Agricultural Frontier of Southern Mexico: A Switching Regression Approach. In: International Regional Science Review, 27(3):326–347.

Wegner, J. (1989): Internationale Märkte für Obst. Analyse von Angebot, Nachfrage, internationalem Handel und Preisen am Beispiel ausgewählter Obstarten. Agrarwirtschaft, Sonderheft 120, Agrimedia, Bergen/Dumme.

Weiss, D. (1995): Preis- und Wechselkurstransmission auf Märkten der Agrar- und Ernährungswirtschaft — Theoretische Analyse und empirische Zusammenhänge am Beispiel des Bananenmarktes. Wissenschaftlicher Fachverlag, Gießen.

Williamson, O. E. (1985): The Economic Institutions of Capitalism. Free Press, New York.

Williamson, O. E. (2000): The New Institutional Economics: Taking Stock, Looking Ahead. In: Journal of Economic Literature, 38:595–613.

Kapitel 2

An Empirical Investigation of the Demand for Bananas in Germany

Alison Burrell und Arne Henningsen

erschienen in: Agrarwirtschaft, Band 50 (2001), S. 242–249.

Abstract

We use econometric methods to investigate consumer demand for bananas and for other fruit in Germany. Monthly household survey data for the period 1986–1998 are analysed. Demand for bananas is significantly responsive to own price, suggesting that policy-induced price increases generate the usual deadweight losses. Demand is also responsive to income changes, indicating that there is scope for further market expansion as incomes grow. There is evidence that other categories of fruit are both gross and net substitutes for bananas.

Keywords: Bananas, Fruit, Dynamic Demand System, Germany

Eine empirische Untersuchung der Bananennachfrage in Deutschland

Wir verwenden ökonometrische Methoden, um die Nachfrage der Konsumenten nach Bananen und anderen Obstarten in Deutschland zu untersuchen. Dabei werden Daten von monatlichen Haushaltsstichproben aus den Jahren 1986–1998 verwendet. Die Bananennachfrage wird sowohl durch den Preis als auch durch die Einkommenshöhe signifikant beeinflusst. Dies lässt die üblichen Wohlfahrtsverluste durch politikbedingte Preisanstiege erwarten und zeigt, dass durch steigendes Einkommen weiteres Potential zur Marktexpansion besteht. Weiter wurde nachgewiesen, dass andere Kategorien von Obst sowohl Brutto- als auch Nettosubstitute für Bananen sind.

Schlüsselwörter: Bananen, Obst, Nachfrage, dynamisches Nachfragesystem, Deutschland

2.1 Introduction

This paper investigates consumer demand for bananas in Germany. We present estimates of the main parameters of interest to market analysts and policy makers. As part of the analysis, demand for other types of fruit is also covered. The analysis uses monthly household survey data for the period 1986–1998. To our knowledge, the agricultural economics literature does not offer any study of demand for bananas in Germany based on recent data, nor any analysis of German consumers' demand for fruit at household level.

Section 2.2 of the paper presents the motivation for our research, describes the policy context and summarises previous research. Section 2.3 describes the data used, and the two modelling approaches adopted. Section 2.4 presents the results, and conclusions are drawn in section 2.5. Additional technical material is available in the appendix.

2.2 Background

In 1997, Germany accounted for 10 per cent of the world's banana imports, representing about 1.1 million tons of the 3.15 million tons of bananas imported into the EU (FAO, 1999). Since the mid-1980s, Germany's per capita banana consumption has been consistently among the highest in Europe. In the last few years, per capita consumption of bananas in Germany was about 13 kilograms per year. This represents a small decline compared with 1991–92, when lower prices and the aftermath of reunification boosted per capita national consumption to record levels of over 15 kg per head.

Most previous studies of the German banana market have focussed on the supply side of the market. Vertical integration in the chain between importation and the retail market is well developed, and there is strong market concentration at the level of importers and ripeners, where three firms share about three quarters of the market. These structural characteristics have motivated research into the functioning of the market. Deodhar and Sheldon (1995) showed that the German banana market is not perfectly competitive, with firms exhibiting Cournot-Nash behaviour. Herrmann and Sexton (1999), however, have argued that weekly import price formation is guided by the Chiquita price. By contrast, Weiss (1995) found strong price transmission from world market price to consumer price, suggesting that there is strong competition despite market concentration.

The German market has also been of interest because, prior to the common banana regime that came into force on 1 July 1993 (Official Journal of the EC, 1993), Germany was the only EU country to enjoy virtually free trade in bananas[1]. Completion of the EU's single internal market in 1993 required the adoption of a common market regulation for bananas[2]. As a consequence, German banana imports from traditional suppliers faced an import quota and a new tariff of 100 ECUs per ton. Kersten (2000) estimated that this policy increased the German wholesale price by around 400 US$/t. We note that German consumers still

[1]Under a special protocol to the Treaty of Rome, Germany benefited from a duty-free quota for banana imports that was nearly sufficient for domestic requirements. This quota was largely filled by imports from Latin America, under the control of a small number of US trading companies. Other EU member countries applied the common external tariff, or imported more expensively produced bananas from ACP countries and overseas territories of Spain, France and Portugal.

[2]The main features of this regulation were a duty-free quota for ACP bananas of 857,700 tons, a tariff rate quota (TRQ) (within-quota tariff=100 ECU/t) of 2 million tons and deficiency payments for EU producers (including overseas territories) up to 854,000 tons. The normal tariff rate on non-quota imports is 750 (850) ECU/t for ACP (non-ACP) countries. For more details, see Kersten (1994); Tangermann (1997); Hallam and Peston (1997). Almost immediately, this regulation was challenged by four Latin American countries under the GATT. As a result, the TRQ was increased progressively by 0.2 mn t and the in-quota tariff was reduced to 75 ECU/t. With the accession of three new member countries in 1995, the TRQ was extended by a further 0.353 mn t. Following further protests, the EU revised its rules for allocating quotas to importing countries.

enjoy the lowest banana prices within the EU (for example, in 1998, retail banana prices in Germany were 11, 10, 16 and 32 per cent lower than in France, the United Kingdom, Italy and Denmark respectively) (FAO, 1999). It is unlikely that the new EU banana regime starting on 1 April 2001 will bring downward pressure on prices[3].

The implications of the 1993 policy change in terms of consumer welfare losses and the adjustment strategies of the market have been studied by Kersten (1995), Herrmann (1999), and Herrmann and Sexton (1999). Despite this recent research, the economic characteristics of the German consumer's demand for bananas have been insufficiently explored. The literature offers conflicting evidence on the responsiveness of consumer demand to changes in banana price, and on the possible existence of close substitutes. Table 2.1 reports empirical estimates taken from recent literature.

All models summarised in Table 2.1 are estimated using annual data. Only Wegner's model is dynamic (short-run elasticities are reported in Table 2.1). Only two of these studies analyse demand at consumer level. Of these, Deodhar and Sheldon (1995) find that consumers have a significant negative own-price response, whereas in Weiss's model the elasticity is significantly different from zero at 10% only. Weiss (1995) finds that bananas and apples are gross substitutes at consumer level, whereas Herrmann and Sexton (1999) argue against any substitution relationship with apples on the basis of their finding that banana imports are not significantly affected by changes in the import price of apples.

The models and estimates of consumer demand behaviour presented in this paper not only complement the studies reported in Table 2.1, but also go beyond them in several respects. First, we extend the time period used for estimation to the end of 1998. Previous studies use data up to 1992 only, on the grounds that subsequent national import or consumption data are less accurate due to the removal of border controls after the advent of the single market. In addition, as far as studies of import demand are concerned, the possibility of structural changes due to the common market regulation introduced in 1993 may have discouraged extending these models beyond 1992. However, for modelling demand at household level, consistently accurate data are available from household surveys right up to the present period. Moreover, it is unlikely that the policy changes introduced during the 1990s had any effect on consumer demand other than via price changes. Therefore, we expect the structure of household demand relationships to have remained unaffected by the policy change.

[3]In December 2000, the Council of Ministers agreed on a new regime comprising three tariff-rate quotas (quota A=2.2 mn t, quota B=0.353 mn t (adjustable) and quota C=0.850 mn t with tariff rates of €75, €75 and €300/t respectively), which will operate for five years from 1 April 2001. These quotas will be open to imports from all third countries. ACP countries will benefit from a tariff preference of €300/t (Official Journal of the EC, 2001)

Table 2.1: Estimated elasticities of demand for bananas from previous studies

Source	Period	Level	Elasticity with respect to price of			Static (S) or dynamic (D)
			bananas	apples	oranges	
WEGNER (1989)[1]	1970–85	Import	-0.29	—	1.06**	D
DEODHAR and SHELDON (1995)[2]	1970–92	Consumer	-0.32**	—	—	S
WEISS (1995)	1970–92	Consumer	-0.42	0.35*	—	S
WEISS (1995)	1970–92	Import	-0.34	0.33*	—	S
HERRMANN (1996)[3]	1960–92	Import	-0.36 / -0.41**	0.19	—	S
HERRMANN and SEXTON (1999)[4]	1977–92	Import	-0.55**	n.s.	—	S

* (**): significant at the 5% (1%) significance level. n.s.: not significant.

[1] Derived from WEGNER (1989)'s semi-logarithmic model (p. 268) using reported average real price (p. 309).

[2] DEODHAR and SHELDON (1995) estimate a linear model (p. 344). The figure shown is the response coefficient $\partial q/\partial p$. Sample averages are not given, thus derivation of an elasticity is not possible.

[3] HERRMANN (1996) finds elasticities in the range -0.36 to -0.41, depending on the model specification.

[4] HERRMANN and SEXTON (1999, p. 13, equation (3)) estimate a linear model, from which they calculate the own-price elasticity (at sample means) shown above. The coefficient on apple price ($-0.4164 \cdot 10^{-05}$) is insignificant, and the corresponding cross-price elasticity for apples is not calculated.

Second, we estimate demand on a monthly basis, thereby allowing a more accurate representation of the links between changes in economic variables and demand. The market for fresh fruit in Germany is highly seasonal. Banana prices are lowest from August to January, apples are cheapest from October to February, and oranges are less expensive from December to March. Clearly, much information is lost in models that seek to capture consumers' reactions to prices using annual data. Contrary to expectations, however, seasonal fluctuations in demand for individual fruits are not simply the inverse of own-price movements. Seasonal fluctuations in banana demand are relatively small (compared to other fresh fruit), with demand highest from March to May and lowest in December, and demand shows a positive correlation with the seasonal variation in banana price. Apple demand has a strong seasonal peak in October, increases again from March to May and slumps between June and September. Demand for oranges is high from December to March, and particularly low from June to September. Modelling demand on a monthly basis, therefore, requires some action to control for seasonal changes in demand that are unrelated to price and income changes. Moreover, since consumers' reactions to a price change may not be completed within a month due to habit persistence (see, for example Alessie and Kapteyn, 1991), dynamic adjustment effects should be included in a monthly demand model.

Third, our data set allows us to complement the aggregate analyses reported above by studying the demand behaviour of several different types of household that vary in terms of composition and income level.

Two sets of model results are presented. First, we report single-equation models that are more directly comparable with the results shown in Table 2.1. Second, we present the results of estimating a demand system for fresh fruit in which bananas and other fruit categories are included. This allows us to account in a systematic way for the possibility of substitutes or complements among different categories of fruit, including bananas.

Both approaches are used to model per capita household demand for each of three clearly defined household types. A description of the data used is given in the following section.

2.3 Data and specification of models

2.3.1 Data

We use data obtained from the German Federal Statistical Office (Statistisches Bundesamt, 1999). The data are derived from sample surveys of households in the former Federal Republic of Germany. Households from the former German Democratic Republic are not included in this study because relevant data are only available for a much shorter period, and because we assume that initial differences between consumption patterns in these two segments will rapidly disappear. Whilst in a strict statistical sense the sample households represent just 5 per cent of all households in the old Länder, their behaviour is most likely typical of large segments of the population from which they are drawn.

The households surveyed are divided into three groups according to structural and income criteria. The income boundaries are adjusted annually in line with the average change in German wages. Households remain in the survey as long as they meet the survey requirements, and are replaced when this is no longer the case. The characteristics and sample averages for the three household types are shown in the Appendix (Table 2.A1). Household type 1 consists largely of older couples whose income is composed mainly of pensions or social security payments. Household type 2 corresponds more or less to the "average" German household, with an average gross income similar to that of a male industrial worker. Household type 3 has the same structure as household type 2, but receives a much higher income. Household type 1 has the highest per capita consumption of all types of fruit, reflecting the fact that these households consist only of adults but suggesting also the possibility that older people consume more fruit for health reasons.

The banana prices paid by the sample households are consistently lower than the price calculated by the Statistisches Bundesamt on the basis of its consumer price survey,[4] although there are strong correlations between all these prices. Statistical tests confirm that the "richer" households (type 3) pay significantly higher prices for bananas than the other two household types, with an average price difference of over 5 per cent.

During the sample period, real expenditure on food and on all fresh fruit, and their shares in total expenditure, declined for all three household types. Taking all households together, the food expenditure share fell from 17.5% to 13.5%. However, real expenditure on bananas remained more or less unchanged, ignoring short-run fluctuations, so that its share in fresh fruit expenditure and in food expenditure increased.

2.3.2 Model specifications

Households are assumed to follow a two-stage budgeting process. At the upper level (first stage), expenditure is allocated to the category fresh fruit, as a function of the fresh fruit price index relative to the prices of other goods, and permanent disposable income[5]. Total expenditure is used as a proxy for permanent disposable income, since actual disposable income as measured in the survey is subject to large seasonal fluctuations due to Christmas and holiday bonuses. The first stage relationship between fresh fruit consumption (fresh fruit expenditure divided by the corresponding price index for fresh fruit), and the fresh fruit price index and total expenditure, both deflated by the consumer price index, is modelled using a simple loglinear function with monthly dummy variables and partial adjustment.

At the second stage, the fresh fruit budget is allocated between n categories of fruit. Two different approaches are used to model demand for bananas at the second stage: a single demand equation for bananas, and a demand system from which mutually consistent demand equations for all n categories of fruit are derived. Price elasticities of demand calculated at the lower stage show responses to price, assuming the budget allocated to fresh fruit remains unchanged (=conditional elasticities). Unconditional price elasticities assume that when

[4] Based on random samples of several retailers in 118 communities, collected on the fifteenth day of each month.

[5] This assumption depends in turn on the assumption that fresh fruit as a group is weakly separable in consumers' utility functions, that is, that consumers' preference orderings between different fruits within the category, conditional upon expenditure on the category as a whole, are independent of the level of consumption of goods outside that category (see Deaton and Muellbauer, 1980a, p. 127)

the price of a fruit category changes, the budget allocated to fresh fruit also adjusts but total expenditure (permanent disposable income) is held constant. Conditional expenditure elasticities show the reaction of demand for a fruit category to a change in the budget allocated to fruit, whereas unconditional expenditure elasticities show the demand response to an increase in total expenditure. Unconditional elasticities are calculated from the results of the two stages of estimation using the formulae given in Fan et al. (1995, p. 62), with standard errors calculated using the formula given in Klein (1953, p. 258).

For the lower level, our first approach involves a loglinear specification with partial adjustment, as shown in (2.1):

$$\log q_{it}^{B} = \beta_{i0} + \sum_{j=1}^{n} \beta_{ij} \log p_{jt}^{r} + \gamma_i \log FX_{it}^{r} + \lambda_i \log q_{it-1}^{B} + \sum_{k=1}^{12} \varphi_{ik} M_{kt} + u_i t, \tag{2.1}$$

for $i = 1, 2, 3$, and for $t = 1, \ldots, 155$, where i denotes household type i, t indexes the months from February 1986 to December 1998, j denotes one of n categories of fruit (where $j = 1$ for bananas), $M_{kt} = 1$ if t falls in the k-th month and zero otherwise, and $\varphi_{i,12} = 0 - \sum_{k=1}^{11} \varphi_{ik}$ to avoid perfect multicollinearity between the monthly dummy variables. q_{it}^{B} is monthly per capita demand for bananas (in grams/head) by household type i, p_{jt}^{r} is the real price index for fruit of category j (1995=100), and FX_{it}^{r} is real expenditure on fresh fruit (in constant 1995 DM) by household type i, all at time t. The consumer price index is used to deflated the price indices and fruit expenditure. In this model, short-run elasticities with respect to prices and fruit expenditure are $\beta_{i1}, \ldots, \beta_{in}, \gamma_i$. Long-run elasticities are $\beta_{i1}/(1 - \lambda_i), \ldots, \beta_{in}/(1 - \lambda_i), \gamma_i/(1 - \lambda_i)$.

For the lower level, our second approach uses a general dynamic version of the linearised Almost Ideal Demand model (Deaton and Muellbauer, 1980b), for which the long-run equilibrium model for the k-th category of fruit is (dropping the subscript i for household types)

$$w_{kt} = \alpha_k + \sum_{j=1}^{n} \alpha_{kj} \log p_{jt} + \beta_k \log \left(FX/P\right)_t \tag{2.2}$$

for $k = 1, \ldots, n$, where w_{kt} is the share of category k in fruit expenditure, p_{jt} is the consumer price index for fruit category j, FX_t is fruit expenditure (in current DM), P_t is a price index for all fruit and the other variables are as before. To obtain a model that is linear in parameters, the price index P is represented by Stone's index, defined (in log form) as $\log P_t = \sum_{j=1}^{n} w_{jt} \log p_{jt}$.

Demand theory places the following restrictions on the system given by (2.2): *adding up*: $\sum_{k=1}^{n} \alpha_k = 1$, $\sum_{k=1}^{n} \alpha_{kj} = 0$, $\sum_{k=1}^{n} \beta_k = 0$, (these restrictions ensure that the shares sum to one); *homogeneity*: $\sum_{j=1}^{n} \alpha_{kj} = 0$, (these restrictions ensure that demands are homogeneous of degree zero in prices and income); and *symmetry*: $\alpha_{kj} = \alpha_{jk}$ for all j and k (these restrictions ensure the symmetry of the Slutsky matrix).

The system given by (2.2) is written in matrix form as

$$\underline{w}_t = \Pi(\alpha, \beta)\underline{x}_t \tag{2.3}$$

where $\underline{w}_t$ is a $n \times 1$ vector, $\Pi(\alpha, \beta)$ is a $n \times k$ matrix of the parameters of the long-run system (2.2) and $\underline{x}_t$ is a $k \times 1$ vector of prices and expenditure.

The general dynamic version of the AID model is described in Anderson and Blundell (1982) and has been applied by various other authors (see, for example Anderson and Blundell, 1983, 1984; Kesavan et al., 1993; McGuirk et al., 1995). Short-run dynamic adjustment takes the form of both autoregressive and moving average (partial adjustment) processes. Assuming first-order processes of both kinds, (2.3) becomes

$$(I - \Gamma L)\underline{w}_t = (\Theta_1 + \Theta_2 L)\underline{x}_t \tag{2.4}$$

where I is the $n \times n$ identity matrix, Γ is a $n \times n$ matrix of adjustment coefficients γ_{ij}, L is the (scalar) lag operator, Θ_1 and Θ_2 are $n \times k$ matrices such that $\Pi(\alpha, \beta) = (I - \Gamma)^{-1}(\Theta_1 + \Theta_2)$. Θ_1 contains the short-run response parameters of the dynamic system.

The system given by (2.4) can also be expressed as an error-correction model by writing

$$\Delta\underline{w}_t = \Theta_1 \Delta\underline{x}_t - (I - \Gamma)(\underline{w}_{t-1} - \Pi(\alpha, \beta)\underline{x}_{t-1}) \tag{2.5}$$

To each equation of (2.5) monthly dummies and a time trend are added (see Kesavan et al., 1993); for the k-th equation, this involves adding the terms $\sum_{j=1}^{12} \varphi_{kj} M_{jt}$ and $\mu_k t$, for which "adding up" requires the parameter restrictions $\sum_{k=1}^{n} \varphi_{kj} = 0$ and $\sum_{k=1}^{n} \mu_k = 0$. As in the single-equation model, in order to avoid perfect multicollinearity between the monthly dummies, we impose the restrictions $\sum_{j=1}^{12} \varphi_{kj} = 0$ for each k; this means that the monthly dummies measure deviations of each month from the annual average (rather than from a "base" month).

As Anderson and Blundell (1982) explain, the adjustment coefficients in the matrix Γ remain unidentified unless additional identifying restrictions are imposed. In this paper, we are not interested in the adjustment process but rather in the parameters of the long-run model, since they are directly comparable with the previously estimated annual elasticities reported in Table 2.1. Therefore, following the procedure described by Anderson and Blun-

DELL (1982), we estimate the system in the form (2.5) without restricting the adjustment process.

There are two advantages of estimating the dynamic model in the form given by (2.5). First, it is possible to impose the theoretical restrictions of symmetry and homogeneity on the long-run parameters. We note, however, that this does not guarantee these properties for the short run. Second, we obtain standard errors of the long-run parameters as part of the estimation output. This simplifies the calculation of standard errors for the long-run elasticities.

Long-run price and expenditure elasticities (see CHALFANT, 1987; GREEN and ALSTON, 1990) are calculated using the formulae

$$\varepsilon_{kj} = \frac{1}{w_k}\left(\alpha_{kj} - \beta_k w_j\right) - \delta_{kj} \tag{2.6}$$

where δ_{kj} is the Kronecker δ and

$$\varepsilon_k = 1 + \frac{\beta_k}{w_k} \tag{2.7}$$

The relative merits of our two approaches for the lower level need some discussion. The long-run AID model conforms fully to the static theory of consumer choice, with dynamic adjustment characteristics superimposed. Since the demands for the different fruit categories are estimated together, all uncompensated (expenditure-constant) and compensated (utility-constant) cross-price elasticities for pairs of categories are mutually consistent. By contrast, the loglinear model is simpler to understand and corresponds to the model specifications underlying the estimates in table 2.1. It is easily shown, however, that the loglinear model is incompatible with consumer theory except in the special case of linear Engel curves ($\varepsilon_k = 1 \; \forall \; k$), unit own-price elasticity and zero cross-price elasticities. Of course, the upper-stage model used here with both approaches suffers from this defect. We use it to obtain elasticities that are comparable with those cited above from the literature.

2.4 Results

2.4.1 Single-equation approach

Initially, the price of oranges was included in these equations. However, in that specification, oranges appeared to be net complements for bananas for two household types, which is counter-intuitive. When banana demand was regressed on all prices, including orange price, and total expenditure (rather than fruit expenditure), both Marshallian and Hicksian elasticities of demand for bananas with respect to the price of oranges were insignificant. It was therefore decided to drop the orange price. Oranges were also dropped from the system model on theoretical grounds (see section 2.4.2). The corresponding total fruit expenditure and fruit price index at the upper stage were modified accordingly.

Table 2.2 presents the results of the single-equation approach. Panel A of this table contains the estimated parameters of the conditional uncompensated elasticities. Summary statistics for these regressions are given in note 1 to the table.

The demand for bananas for all three household types responds significantly to own price changes. Long-run responses are much greater than short-run responses, with household type 1 reallocating its expenditure among different types of fruit after a price change most quickly. For household type 1, 96 per cent of the adjustment is complete 2 months after the month of the price change, whereas household type 2 takes 3 months and household type 3 takes 4 months to make 96 per cent of its adjustment.

The compensated elasticities shown in panel C of Table 2.2 suggest that bananas and the category "other fruit" (comprising mainly grapes, peaches, berries, plums, cherries, lemons, grapefruit and tropical fruit (ZMP, 1998, pp. 33, 35, 36)) are weak net substitutes for household type 1. Other relationships of substitutability or complementarity between the different types of fruit were not found. Behaviour is significantly different between the three household types[6].

Elasticities of demand with respect to total expenditure (our proxy for permanent disposable income) indicate that there is still some scope for demand to grow with income for all household types, with the greatest potential exhibited by household type 3. This suggests that, despite high consumption levels in Germany, the market is not yet saturated.

2.4.2 Demand system

The demand system was originally estimated with four categories of fruit: bananas, apples and pears ("apples")[7], oranges and mandarins ("oranges") and the category other fruit ("other") defined as in section 2.4.1. With long-run homogeneity and symmetry imposed, the own-price elasticity for oranges for all three household types violated the necessary condition for concavity, and oranges exhibited significant complementarity with apples, which is counter-intuitive. Imposing concavity at minimum cost to the goodness-of-fit of the system effectively resulted in forcing to zero all elements in the Slutsky matrix corresponding

[6] An F-test rejected a model where parameters were restricted to be the same for all household types in favour of a model where each household type may have different parameters ($F_{36,411} = 7.95$).

[7] An index of the (weighted) prices of these two fruits was used.

Table 2.2: Elasticities of demand for bananas (single equation)[1]

	Household type 1		Household type 2		Household type 3	
	Short run	Long run	Short run	Long run	Short run	Long run
A. Conditional uncompensated elasticities with respect to						
Banana price	– 0.63 **	– 0.91 **	– 0.38 **	– 0.64 **	– 0.26 **	– 0.54 **
Apple price	– 0.06	– 0.08	– 0.13	– 0.21	– 0.11	– 0.22
Other fruit price	0.11	0.16	– 0.09	– 0.15	– 0.11	– 0.22
Fruit expenditure[2]	0.52 **	0.75 **	0.45 **	0.75 **	0.45 **	0.92 **
B. Unconditional[3] uncompensated elasticities with respect to						
Banana price	– 0.53 **	– 0.79 **	– 0.26 **	– 0.45 **	– 0.16 **	– 0.33 **
Apple price	0.10	0.12	0.04	0.07	0.06	0.12
Other fruit price	0.32 +	0.43	0.07	0.11	0.07	0.15
Total expenditure	0.15 **	0.45 **	0.09 +	0.41 +	0.16 *	0.68 *
C. Conditional compensated[4] elasticities with respect to						
Banana price	– 0.52 **	– 0.75 **	– 0.26 **	– 0.44 **	– 0.16 *	– 0.33 **
Apple price	0.12	0.18	0.04	0.07	0.06	0.12
Other fruit price	0.35 +	0.50 +	0.07	0.12	0.07	0.15

+ [*] (**) denotes significantly different from zero in a 2-tailed test at the 10 [5] (1) % significance level.

[1] $\bar{R}^2$ = 0.764, 0.777, 0.815, and Durbin's h (see Johnston and DiNardo, 1997, pp.182–184) = -0.84, -1.29, -4.14 respectively for the three household types. The coefficients on the lagged endogenous variables are 0.31 (4.72), 0.41 (6.40), 0.51 (7.04) respectively (t-ratios in parentheses).

[2] Fruit expenditure does not contain the expenditure on oranges.

[3] The unconditional elasticities are calculated using the results of the upper stage, shown in Table 2.A2.

[4] The compensated elasticities (ε^*_{kj}) are calculated as $\varepsilon^*_{kj} = \varepsilon_{kj} + w_j \varepsilon_k$.

to oranges. Using the test described by Sellen and Goddard (1997), we tested for weak separability between oranges and the other three fruit categories. For household types 2 and 3, the null hypothesis of weak separability was accepted with p-values of over 0.12.[8] For household type 1, the test was inconclusive since the test model failed to converge. Based on this partial evidence in favour of weak separability, we dropped the oranges category from the system for all three household types.

[8] This implies that consumers view oranges and the composite group composed of bananas, apples and other fruit as competing categories at the same budgeting level. Changes in the price of oranges affect demand for bananas only via a change in the amount allocated to be spent on this composite group, ruling out any direct substitution or complementarity.

In the three-fruit system, the share equations for bananas and apples were estimated with long-run homogeneity and symmetry imposed, and the adding-up restrictions on the long-run parameters were used to derive the parameters of the third equation (other fruit)[9]. Concavity was satisfied at the sample means for all three household types. A time trend was included to capture any long-run shifts in preferences between the fruit categories; in the banana equations it was significant only for household type 1, and indicated a ceteris paribus cumulative decline in the banana expenditure share of 1.3 percentage points over the whole sample period for this household type.

The estimated demand system of household type 1 was significantly different from that of household types 2 and 3 at the 1 per cent significance level. By contrast, the estimated systems of household types 2 and 3 were not significantly different from each other at the 10 per cent level[10]. However, we have kept household types 2 and 3 separate because they differ in their upper-level behaviour.

The elasticities obtained from the demand system are summarised in Table 2.3. Panels A and B show long-run conditional and unconditional elasticities respectively. First, we note that demand for bananas responds significantly to changes in own price for all three household types, although banana demand is more inelastic for household types 2 and 3. Demand for the other two fruit categories is also significantly responsive to changes in own price. Demand for other fruit is the most responsive to changes in its own price for household types 2 and 3, whereas demand for bananas is the most price responsive for household type 1.

Second, long-run conditional elasticities of demand for bananas with respect to expenditure are significantly positive for all household types. Long-run unconditional income elasticities for bananas are also significant for all household types, although only weakly significant for household type 2. The overall results suggest that there is potential for growth in demand for all three categories of fruit as per capita income increases. For household types 1 and 3, the potential for income-driven growth in banana consumption is similar to that for apples, but lower than that for other fruit. For household type 2, growth potential is lower for bananas than for the other two categories.

Third, the greater input of data and economic theory in the demand system, relative to the single-equation approach, enables a more detailed investigation of substitution and complementarity relationships between the three categories of fruit. For household type 1,

[9] When homogeneity and symmetry were tested together using a likelihood ratio test, they were rejected with χ^2 (3) values of 28.8, 36.7 and 46.1 respectively (critical value = 7.8). Given our aim to obtain a theoretically consistent set of demand parameters, we give priority to theory. However, we note that these properties have been imposed on the data at some statistical cost.

[10] According to pairwise likelihood ratio tests of pooled models. Chi-square (df=43) values were 77.7, 70.0 and 11.7 respectively.

bananas and apples are long-run gross substitutes in both the banana and the apple demand functions for household type 1, whereas other fruit are gross substitutes for bananas and apples only in the banana and apple equations. For the other two household types, there are two-way relationships of gross substitution between other fruit and bananas, and other fruit and apples, but no gross substitution between bananas and apples. The C panels in Table 2.3 indicate that all three fruit categories are net (Hicksian) substitutes for each other for household type 1, whereas all pairs except bananas and apples are net substitutes for household types 2 and 3.

2.5 Conclusions

This paper presents two different sets of parameter estimates for German households' demand for bananas. Demand patterns have remained quite stable during the sample period, despite underlying policy changes. Bananas are a normal good for German households (i.e. with an income elasticity between zero and one). Clearly, this market is not saturated and has potential for further expansion as incomes rise. According to the single-equation results, households are found to adjust their demands to price changes within 2–4 months after the month of the price change. Demand for bananas is significantly responsive to own price for all household groups, indicating that the 1993 policy change generated the usual deadweight losses.

The range of own-price demand elasticities for bananas is greater in the single equation model (unconditional long-run own-price elasticities between -0.79 for low-income households and -0.33 for high-income households) than in the demand system (unconditional long-run own-price elasticities between -0.60 and -0.42 for low- and medium-income households respectively). Long-run elasticities of demand with respect to total expenditure, which is used here as a proxy for "permanent" disposable income, are also considerably larger in the single-equation model than in the demand system. Price responsiveness is generally lower for the higher-income households in both models.

The higher banana consumption and greater price sensitivity of lower-income households suggests that if the new EU banana regulation that becomes operational on 1 April 2001 were to result in a price reduction for bananas in Germany, the welfare of low-income households, when measured in money terms, would increase relatively more than that of higher-income households. Given that the marginal utility of money is higher for these households, the relative welfare effect in utility terms would of course be even greater.

Table 2.3: Elasticities of demand for three fruit categories

Demand for	Bananas	Apples	Other fruit
Household type 1[1]			
A. Long-run conditional uncompensated elasticities with respect to			
Banana price	– 0.71 **	0.03	– 0.15 **
Apple price	0.03	– 0.64 **	– 0.29 **
Price of other fruit	– 0.04	– 0.06	– 0.94 **
Fruit expenditure	0.72 **	0.68 **	1.37 **
B. Long-run unconditional uncompensated elasticities with respect to			
Banana price	– 0.60 **	0.13 **	0.07
Apple price	0.22 **	– 0.46 **	0.08
Price of other fruit	0.22 *	0.18 +	– 0.45 **
Total expenditure	0.43 **	0.41 *	0.83 **
C. Long-run conditional compensated elasticities (holding utility constant) with respect to			
Banana price	– 0.56 **	0.16 **	0.13 **
Apple price	0.28 **	– 0.41 **	0.19 *
Price of other fruit	0.29 **	0.25 *	– 0.32 **
Household type 2[2]			
A. Long-run conditional uncompensated elasticities with respect to			
Banana price	– 0.59 **	– 0.23 **	– 0.06
Apple price	– 0.15	– 0.76 **	– 0.14
Price of other fruit	0.10	– 0.10	– 0.96 **
Fruit expenditure	0.65 **	1.09 **	1.16 **
B. Long-run unconditional uncompensated elasticities with respect to			
Banana price	– 0.42 **	0.05	0.23 *
Apple price	0.08	– 0.36 *	0.29 +
Price of other fruit	0.32 **	0.27 +	– 0.56 **
Total expenditure	0.35 +	0.60 +	0.64 +
C. Long-run conditional compensated elasticities (holding utility constant) with respect to			
Banana price	– 0.42 **	0.06	0.24 **
Apple price	0.09	– 0.35 *	0.30 *
Price of other fruit	0.33 **	0.28 *	– 0.55 **
Household type 3[3]			
A. Long-run conditional uncompensated elasticities with respect to			
Banana price	– 0.63 **	– 0.16 *	– 0.06
Apple price	– 0.26 +	– 0.69 **	– 0.14
Price of other fruit	0.06	0.01	– 1.04 **
Fruit expenditure	0.83 **	0.84 **	1.24 **
B. Long-run unconditional uncompensated elasticities with respect to			
Banana price	– 0.44 **	0.03	0.23 **
Apple price	0.05	– 0.37 **	0.33 **
Price of other fruit	0.40 **	0.35 **	– 0.54 **
Total expenditure	0.62 *	0.63 *	0.92 *
C. Long-run conditional compensated elasticities (holding utility constant) with respect to			
Banana price	– 0.44 **	0.03	0.22 **
Apple price	0.04	– 0.37 **	0.32 **
Price of other fruit	0.40 **	0.35 **	– 0.54 **

+ [*] (**) denotes significantly different from zero in a 2-tailed test at the 10 [5] (1) % significance level.
[1] $\bar{R}^2$ = 0.830, 0.787, and Durbin Watson = 2.11, 2.18 respectively for the first two equations of panel A.
[2] $\bar{R}^2$ = 0.885, 0.727, and Durbin Watson = 2.28, 2.08 respectively for the first two equations of panel A.
[3] $\bar{R}^2$ = 0.906, 0.806, and Durbin Watson = 2.10, 2.10 respectively for the first two equations of panel A.

The system estimates show that bananas are net substitutes for apples and pears and for other fruit for low-income households, and for other fruit only for the other two household types. This pattern is largely maintained when the income effects of price changes are allowed for: in the system-estimated banana demand function, bananas and the category of other fruit are gross substitutes for all household types and in addition, bananas and apples are gross substitutes for the lowest-income households. This result contradicts the statements from the literature referred to above that suggest bananas have no close substitutes for the German consumer. We note, however, that bananas do not appear to have gross substitutes in the single-equation model.

Although our results are based on data for a small, clearly defined proportion of German households, we consider that the behaviour described here is probably typical of a large segment of the German market. It remains a disadvantage, of course, that our estimates do not directly describe the aggregate behaviour of the German market. The advantages of using household survey information, however, are important. They include the high quality and consistency of the data, the fact that demand behaviour can be analysed as close as possible to the decision-making unit (as opposed to market data, which do not always coincide exactly with what households have purchased), the ability to obtain up-to-date estimates, and the possibility of identifying different responses at different levels in the income distribution.

Appendix

2.6 Data description

Table 2.A1: Description of households and household data

	Household type 1	Household type 2	Household type 3
Household characteristics			
Approximate sample size	165	378	388
Composition	2 adults	2 adults, 2 children	2 adults, 2 children
Age of children	—	at least one < 15	at least one < 15
Employment status	mainly not working	only 1 adult working	at least 1 adult working
Average values (1986-1998) in DM (1995)/month			
Total disposable income	2,555	5,313	8,626
Total consumption expenditure	2,129	4,082	5,991
Total food expenditure	394	610	738
Per capita expenditure on fresh fruit	14.93	8.41	11.32
bananas	2.39	1.78	2.07
apples	4.03	2.68	3.53
other fruit	6.13	2.75	4.18
oranges	2.38	1.20	1.54
Average expenditure shares (1986-1998) in per cent			
% Share food in total consumption	18.5	14.9	12.3
% Share of fresh fruit (including oranges) in total food	7.6	5.5	6.1
% Share in expenditure on fresh fruit (excluding oranges) of			
bananas	19.0	24.7	21.2
apples	32.1	37.2	36.1
other fruit	48.8	38.1	42.7

Source: Own computations with data from the German Federal Statistical Office (Statistisches Bundesamt)

2.7 Results of upper-stage estimation

Table 2.A2: First stage: uncompensated elasticities of fruit demand[1]

	Household type 1		Household type 2		Household type 3	
	Elasticity	t-ratio	Elasticity	t-ratio	Elasticity	t-ratio
Short-run elasticities with respect to						
Price of fruit	– 0.10	– 1.37	– 0.01	– 0.15	0.01	0.11
Total expenditure	0.28	2.96	0.20	1.86	0.36	2.67
Long-run elasticities with respect to						
Price of fruit	– 0.22	– 1.34	– 0.03	– 0.15	0.02	0.11
Total expenditure	0.60	3.28	0.55	1.86	0.74	2.46

[1] Fruit demand and the fruit price index do not contain the expenditure on oranges or the orange price respectively. $\bar{R}^2$ = 0.923, 0.933, 0.929. Durbin's h (see Johnston and DiNardo, 1997, pp. 182–184) = -0.810, -0.544, 1.858 respectively for the three household types. The coefficients on the lagged endogenous variables are 0.54 (7.68), 0.64 (8.76), 0.51 (6.69) respectively (t-ratios in parentheses).

2.8 Calculation of unconditional elasticities (Fan, Wailes and Cramer (1995))

2.8.1 Unconditional price elasticities

$$E_{ij}^{\bar{X}} = E_{ij}^{\bar{E}} + E_{iE} w_j (E_{FP}^{\bar{X}} + 1) \tag{2.8}$$

where:

$E_{ij}^{\bar{X}}$ = Elasticity of demand for fruit i with respect to the price of fruit j holding total expenditure (X) constant

$E_{ij}^{\bar{E}}$ = Elasticity of demand for fruit i with respect to the price of fruit j holding expenditure on fruit (E) constant

E_{iE} = Elasticity of demand for fruit i with respect to expenditure on fruit

w_j = Expenditure share of fruit j of expenditure on all fruit

$E_{FP}^{\bar{X}}$ = Elasticity of demand for total fruit with respect to fruit price (P) holding total expenditure constant

2.8.2 Unconditional expenditure elasticities

$$E_{iX} = E_{iE} E_{FX} \tag{2.9}$$

where

E_{iX} = Elasticity of demand for fruit i with respect to total expenditure

E_{iE} = Elasticity of demand for fruit i with respect to expenditure on fruit

E_{FX} = Elasticity of total fruit demand with respect to total expenditure

Bibliography

ALESSIE, R. and A. KAPTEYN (1991): Habit Formation, Interdependent Preferences and Demographic Effects in the Almost Ideal Demand System. In: The Economic Journal, 101:404–419.

ANDERSON, G. J. and R. W. BLUNDELL (1982): Estimation and Hypothesis Testing in Dynamic Singular Equation Systems. In: Econometrica, 50(6):1559–1571.

ANDERSON, G. J. and R. W. BLUNDELL (1983): Testing Restrictions in a Flexible Dynamic Demand System: An Application to Consumers' Expenditure in Canada. In: Review of Economic Studies, 50:397–410.

ANDERSON, G. J. and R. W. BLUNDELL (1984): Consumer non-durables in the U.K.: A Dynamic Demand System. In: The Economic Journal, 94 (Supplement):35–44.

CHALFANT, J. A. (1987): A Globally Flexible, Almost Ideal Demand System. In: Journal of Business and Economic Statistics, 5:233–242.

DEATON, A. and J. MUELLBAUER (1980a): An Almost Ideal Demand System. In: The American Economic Review, 70:312–326.

DEATON, A. and J. MUELLBAUER (1980b): Economics and Consumer Behaviour. Cambridge University Press, Cambridge.

DEODHAR, S. Y. and I. SHELDON (1995): Is Foreign Trade (Im)Perfectly Competitive?: An Analysis of the German Market for Banana Imports. In: Journal of Agricultural Economics, 46(3):336–348.

Fan, S., E. J. Wailes and G. L. Cramer (1995): Household Demand in Rural China: A Two-Stage LES-AIDS Model. In: American Journal of Agricultural Economics, 77:54–62.

FAO (1999): Banana information note. ESC; BA/99, December 1999.

Green, R. and J. M. Alston (1990): Elasticities in AIDS models. In: American Journal of Agricultural Economics, 72:442–445.

Hallam, D. and L. Peston (1997): The Political Economy of Europe's Banana Trade. Occasional Paper 5, Department of Agricultural and Food Economics, University of Reading, Reading.

Herrmann, R. (1996): Economic Implications of the New European Banana Market Regime: The Case of Germany. Discussion Papers in Agricultural Economics 38, Department of Agricultural Economics, University of Gießen, Gießen.

Herrmann, R. (1999): Economic Impacts of the New European Banana Market Regime: The Case of Germany. In: Jahrbücher für Nationalökonomie und Statistik, 218(1+2):63–84.

Herrmann, R. and R. Sexton (1999): Redistributive Implications of a Tariff-rate Quota Policy: How Market Structure and Conduct Matter. Discussion Papers in Agricultural Economics 52, Department of Agricultural Economics, University of Gießen, Gießen.

Johnston, J. and J. DiNardo (1997): Econometric Methods. 4th edition. McGraw-Hill.

Kersten, L. (1994): Wirkungen der EWG-Marktordnung für Bananen auf Handel und Preisbildung. In: Agrarwirtschaft, 43(8/9):319–327.

Kersten, L. (1995): Impacts of the EU Banana Market Regulation on International Competition, Trade and Welfare. In: European Review of Agricultural Economics, 22:321–335.

Kersten, L. (2000): Zur Revision der Bananenmarktordnung. Unveröffentlicht. Zusammenfassung veröffentlicht im Jahresbericht 1999 der FAL Braunschweig-Völkenrode, Kapitel 5.1, S. 136-137.

Kesavan, T., Z. A. Hassan, H. H. Jensen and S. R. Johnson (1993): Dynamics and Long-run Structure in U.S. Meat Demand. In: Canadian Journal of Agricultural Economics, 41:139–153.

Klein, L. R. (1953): A Textbook of Econometrics. Row, Peterson and Co, New York.

McGuirk, A., P. J. Driscoll, J. Alwang and H. Huang (1995): System Misspecification Testing and Structural Change in the Demand for Meats. In: Journal of Agricultural and Land Resource Economics, 20(1):1–21.

Official Journal of the EC (1993): Council Regulation (EEC) No 404/93 of 13 February 1993 on the Common Organisation of the Market in Bananas. L 047, 25/02/1993, p. 1-11.

Official Journal of the EC (2001): Council Regulation (EC) No 216/2001 of 29 January 2001 on the Common Organisation of the Market in Bananas. L 031, 2/02/2001, p. 2-4.

Sellen, D. and E. Goddard (1997): Weak Separability in Coffe Demand Systems. In: European Review of Agricultural Economics, 24(1):133–144.

Statistisches Bundesamt (1999): Zeitreihenservice. http://www-zr.statistik-bund.de/home.htm, various time series.

Tangermann, S. (1997): Banana Policy: A European Perspective. In: Australian Journal of Agricultural and Resource Economics, 41(2):263–267.

Wegner, J. (1989): Internationale Märkte für Obst. Analyse von Angebot, Nachfrage, internationalem Handel und Preisen am Beispiel ausgewählter Obstarten. Agrarwirtschaft, Sonderheft 120, Agrimedia, Bergen/Dumme.

Weiss, D. (1995): Preis- und Wechselkurstransmission auf Märkten der Agrar- und Ernährungswirtschaft — Theoretische Analyse und empirische Zusammenhänge am Beispiel des Bananenmarktes. Wissenschaftlicher Fachverlag, Gießen.

ZMP (1998): ZMP-Bilanz Obst. Zentrale Markt- und Preisberichtstelle für Erzeugnisse der Land-, Forst- und Ernährungswirtschaft, Bonn-Bad Godesberg.

Kapitel 3

Die EU Bananenmarktordnung und die Nachfrage in Deutschland

Arne Henningsen

erschienen in: Penker, M. und Pfusterschmid, S. (Hrsg.) (2003): Wie steuerbar ist die Landwirtschaft? Erfordernisse, Potentiale und Instrumente zur Ökologisierung der Landwirtschaft. Dokumentation der 11. ÖGA-Jahrestagung an der Karl-Franzens-Universität Graz, 27. und 28. September 2001, Facultas, Wien, S. 157–164.

Zusammenfassung

Die deutschen Verbraucher waren durch die Einführung der EU Bananenmarktordnung am stärksten betroffen. Mit Hilfe der Ergebnisse der Nachfrageanalyse von Burrell und Henningsen (2001) werden die Auswirkungen auf die Frischobstnachfrage deutscher Haushalte ermittelt. Wegen des durch die Marktordnung bedingten Preisanstiegs sinkt der Bananenkonsum in den untersuchten Haushaltstypen zwischen 17,6% und 12,5%, während aufgrund von Substitutionsbeziehungen der Konsum von Kernobst sowie der Kategorie „anderes Frischobst“ ansteigt. Die Ausgaben der Haushalte für Frischobst insgesamt steigen zwischen 4,7% und 6,1%.

Schlagworte: EU Bananenmarktordnung, Nachfrage, Deutschland.

Summary

The German consumers have been hurt most by the implementation of the EU banana regime. Its impacts on the demand for fresh fruit of German households is analysed by using the results of the demand analysis by Burrell und Henningsen (2001). The policy-induced price increase of bananas has reduced the banana consumption of the examined household types by 17.6% to 12.5% and increased the consumption of substitutes, namely apples and pears and the category "other types of fresh fruit". The households' expenditure on total fresh fruit increased by 4.7% to 6.1%.

Keywords: EU banana regime, demand, Germany.

3.1 Einleitung

Einer der größten Streitpunkte innerhalb der EU sowie zwischen der EU und ihren Handelspartnern war die Gemeinsame Marktordnung für Bananen. Bei ihrer Einführung im Jahr 1993 waren die deutschen Verbraucher am stärksten betroffen, denn zum einen hatte Deutschland zuvor durch zollfreie Importe die niedrigsten Bananenpreise in der EU und musste somit den stärksten Preisanstieg hinnehmen und zum anderen ist der Pro-Kopf-Verbrauch von Bananen in Deutschland einer der höchsten in der EU. Die Auswirkungen der EU Bananenmarktordnung auf den Welthandel wurden vielfach untersucht. Als Ergänzung dazu werden in diesem Beitrag die Auswirkungen auf verschiedene deutsche Haushaltstypen ermittelt.

3.2 Die Europäische Bananenpolitik

In diesem Artikel kann nur auf die wichtigsten Elemente der europäischen Bananenpolitik eingegangen werden. Weitere Einzelheiten können z.B. bei Hallam und Peston (1997) und aktuelle Entwicklungen auf den Internetseiten der EU (http://www.europa.eu.int) nachgelesen werden.

3.2.1 Die nationalen Politiken vor 1993

Vor Einführung der Gemeinsamen Marktordung für Bananen im Jahr 1993 gab es in der EU eine Vielzahl von verschiedenen nationalen Marktregelungen, die allerdings in zwei Gruppen eingeteilt werden konnten: Länder mit relativ freien und Länder mit geschützten Märkten. Zu den ersteren zählten Belgien, Dänemark, Deutschland, Irland, Luxemburg und die Niederlande, deren Bananenimporte keinen quantitativen Beschränkungen unterlagen. Während Bananen zollfrei nach Deutschland importiert werden konnten, unterlagen Importe in jene andere Länder einem 20%igen Wertzoll. Diese Märkte wurden fast ausschließlich mit Bananen aus Lateinamerika versorgt. Zur zweiten Gruppe gehörten Frankreich, Griechenland, Italien, Portugal, Spanien und das Vereinigte Königreich, die ihre Märkte hauptsächlich für die inländische Produktion (Kreta, Kanarische Inseln, Madeira, Martinique und Guadeloupe) und für Importe aus assoziierten Staaten (ehemaligen Kolonien und Commonwealth-Staaten) reservierten.

3.2.2 Die Gemeinsame Marktordnung für Bananen

Die Einführung des Europäischen Binnenmarktes zum 1.1.1993 bedeutete die Aufgabe sämtlicher innergemeinschaftlicher Handelsbarrieren. Daher mussten auch die verschiedenen nationalen Regelungen des Bananenmarktes durch eine Gemeinsame Bananenmarktordnung ersetzt werden. Diese wurde wegen der unterschiedlichen Interessenlagen der Mitgliedsländer lange und kontrovers diskutiert und schließlich im Februar 1993 beschlossen. Sie enthielt Preisausgleichzahlungen für EU-Produzenten und getrennte Importquoten für traditionelle Einfuhren aus AKP (= Afrika, Karibik, Pazifik) -Staaten einerseits (857.700 t) sowie für nicht-traditionelle Einfuhren aus AKP-Staaten und Importe aus Drittländern andererseits (2 Mio. t), wobei die Einfuhren aus Drittländern zusätzlich noch einem Zoll unterlagen.

3.2.3 Änderungen der Gemeinsamen Bananenmarktordnung

Nach einer erfolgreichen Beschwerde von fünf lateinamerikanischen Staaten vor dem GATT-Panel im Januar 1994 einigte sich die EU mit vier dieser Staaten auf eine Änderung der Bananenmarktordnung. Im Zuge dieses sogenannten Rahmenabkommens wurde die Quote für Drittlands- und nicht-traditionelle AKP-Bananen auf 2,2 Mio. t erhöht und der Zollsatz verringert. Weiterhin wurde im Zuge der EU-Erweiterung 1995 die jährliche Quote für Drittlands- und nicht-traditionelle AKP-Bananen um 353.000 t auf nun 2,553 Mio. t erhöht. Doch der Streit um die EU-Bananenmarktordnung war noch nicht beigelegt, denn im September 1997 entschied das Streitschlichtungsgremium der WTO aufgrund einer Beschwerde der USA und vier lateinamerikanischen Staaten, dass auch die veränderte Bananenmarktordnung nicht den GATT-Prinzipien entspräche. Daraufhin änderte die EU das Verfahren zur Verteilung der Importlizenzen zugunsten US-amerikanischer Handelsunternehmen und lateinamerikanischer Exporteure. Während die EU-Kommission die so veränderte Bananenmarktordnung für WTO-konform hielt, befand das Streitschlichtungsgremien der WTO auch in diesem Fall, dass einige Elemente der EU-Bananenmarktordnung den Prinzipien des GATT widersprächen, und erlaubte deshalb den USA, Strafzölle auf bestimmte EU-Importe zu erheben. Nach jahrelangem Streit erreichte die EU schließlich im April 2001 eine Einigung mit den USA und Ecuador. Im Rahmen dieser Einigung verpflichtete sich die EU, spätestens zum 1. Januar 2006 ein reines Zollsystem für Bananenimporte einzuführen und eine Übergangsregelung mit weiteren Verbesserungen für lateinamerikanische Exporteure und US-amerikanische Handelsfirmen zu implementieren. Im Gegenzug verpflichtete sich die USA, die Strafzölle einzustellen. Diese Übergangsregelung trat zum 1. Juli 2001 in Kraft.

3.3 Die Bananennachfrage in Deutschland

Um die Auswirkungen der europäischen Bananenmarktordnung auf die deutschen Verbraucher näher zu analysieren, sind detaillierte Untersuchungen der Bananennachfrage notwendig. Dies soll auf Grundlage eigener Untersuchungen (siehe Burrell und Henningsen, 2001) geschehen. Da diese Untersuchungen bereits in dem genannten Artikel beschrieben sind, werden hier nur die für die weitere Analyse wichtigsten Informationen dargestellt.

3.3.1 Modellspezifikationen

Es wird angenommen, dass die Haushalte ihr Budget in einem zweistufigen Prozess aufteilen. Auf der ersten Stufe teilen sie einen bestimmten Anteil ihres Einkommens den Ausgaben für

Frischobst zu, der vom Frischobstpreisindex und den gesamten Konsumausgaben abhängt. Diese Nachfragefunktion für Frischobst wurde mit einem dynamischen log-linearen Modell ökonometrisch geschätzt.

Auf der zweiten Stufe werden die Ausgaben für Frischobst auf drei verschiedenen Frischobstkategorien (Bananen, Kernobst und anderes Frischobst) aufgeteilt. Dieser Prozess wurde mit einem Nachfragesystem untersucht und in einer allgemeinen dynamischen Form des linearisierten „Almost Ideal Demand Systems“ (AIDS) (siehe Deaton und Muellbauer, 1980; Anderson und Blundell, 1983) geschätzt.

3.3.2 Datengrundlage

Beide Modelle basieren auf monatlichen Preisen und Haushaltsausgaben der Jahre 1986–1998. Es wurden Preise (bzw. Preisindizes) und Haushaltsausgaben für die drei Frischobstkategorien, für Frischobst insgesamt und den gesamten Konsum verwendet, die vom Statistischen Bundesamt durch Preisbeobachtungen in Einzelhandelsgeschäften bzw. aus Haushaltsstichproben ermittelt wurden. Dabei werden drei Haushaltstypen unterschieden: Haushaltstyp 1 sind 2-Personen-Haushalte, die vorwiegend aus älteren Personen mit niedrigem Einkommen bestehen. Typ 2 sind 4-Personen-Haushalte, die aus zwei Erwachsenen und zwei Kindern bestehen und ein Haushaltseinkommen in Höhe des durchschnittlichen Lohnes eines deutschen Arbeiters haben. Typ 3 hat die gleiche Struktur wie Typ 2 nur ein wesentlich höheres Haushaltseinkommen.

3.3.3 Ergebnisse

Aus den Regressionsergebnissen der Nachfragefunktion für Frischobst (erste Stufe) wurden die Elastizitäten der Frischobstnachrage ermittelt. Die „unbedingten“ (d.h. nicht an die Bedingung konstanter Frischobstausgaben gebundene, siehe Fan et al., 1995) Nachfrageelastizitäten der einzelnen Frischobstkategorien wurden anschließend aus den Regressionsergebnissen beider Stufen berechnet (siehe Tabelle 3.1).

Die Gesamtmenge des nachgefragten Frischobstes steigt mit zunehmenden Einkommen, hängt aber nicht (signifikant) vom Preisindex für Frischobst ab. Dagegen wird die Zusammensetzung des nachgefragten Frischobstes durch eine Änderung des Preises nur einer Frischobstkategorie signifikant beeinflusst. Dabei ist die Bananennachfrage des „ärmeren“ Haushaltstyps 1 elastischer als die der anderen Haushaltstypen. Kernobst (Haushaltstyp 1) bzw. anderes Frischobst (Typ 2 u. 3) werden von den Haushalten als Substitute für Bananen betrachtet. Die Verbraucher zeigen beim Frischobstkauf deutlich Gewohnheitsverhalten, was

Tabelle 3.1: Ergebnisse der Nachfrageanalyse

	Haushaltstyp 1		Haushaltstyp 2		Haushaltstyp 3	
	kurzfr.	langfr.	kurzfr.	langfr.	kurzfr.	langfr.
Elastizitäten der Frischobstnachfrage in Bezug auf						
Frischobstpreis	– 0,10	– 0,22	– 0,01	– 0,03	0,01	0,02
Gesamtausgaben	0,28 **	0,60 **	0,20 +	0,55 +	0,36 **	0,74 *
Elastizitäten Bananennachfrage in Bezug auf						
Bananenpreis	– 0,44 **	– 0,60 **	– 0,28 **	– 0,42 **	– 0,20 **	– 0,44 **
Kernobstpreis	0,36 +	0,22 **	0,08	0,08	0,15	0,05
Preis von anderem Frischobst	0,26	0,22 *	0,29 +	0,32 **	0,26	0,40 **
Gesamtausgaben	0,14 **	0,43 **	0,11 +	0,35 *	0,21 *	0,62 *
Elastizitäten der Nachfrage nach ... in Bezug auf den Bananenpreis						
Kernobst	0,20 +	0,13 **	0,05	0,05	0,01	0,03
And. Frischobst	0,00	0,07	0,14 **	0,23 *	0,11 *	0,23 **

+, * und ** bedeuten ein Signifikanzniveau von 10%, 5% bzw. 1%.

Quelle: Burrell und Henningsen (2001), eigene Berechnungen

dadurch zum Ausdruck kommt, dass die kurzfristige Nachfrage unelastischer als die langfristige ist.

3.4 Die Auswirkungen der Bananenmarktordnung

3.4.1 Auswirkungen auf Verbraucherpreise

Die ökonomischen Auswirkungen der Bananenmarktordnung sind in zahlreichen Studien analysiert worden. Durch ihre Einführung sind aufgrund der Quotierung der Bananenimporte die Preise auf den vormals relativ freien Märkten deutlich gestiegen. Die nach Wissen des Autors einzigen Studien, die explizit die Auswirkungen auf die deutschen Verbraucherpreise angeben, sind dabei die von Kersten (z.B. 1995, 2000). Darum beruhen folgende Berechnungen auf dessen Ergebnissen. Auf Grundlage der Ergebnisse von Kersten (2000) ergaben eigene Berechnungen, dass die deutschen Verbraucherpreise ca. 38% über dem hypothetischen Preis bei einer vollkommenen Liberalisierung liegen. Dieses Ergebnis bezieht sich auf die Bananenmarktordnung in der Form nach dem Rahmenabkommen und der EU-Erweiterung 1995. Es wird angenommen, dass die weiteren Änderungen der Bananenmarktordnung keinen Einfluss auf die Verbraucherpreise hatten, da die gesamte Importquote nicht verändert wurde.

Tabelle 3.2: Auswirkungen der Bananenmarktordnung

	Haushaltstyp 1	Haushaltstyp 2	Haushaltstyp 3
Änderungen der Nachfragemengen [%] von ...			
Frischobst (Index)	– 1,3	– 0,2	0,1
Bananen	– 17,6	– 13,2	– 12,5
Kernobst	3,8	– 0,2	0,2
Anderes Frischobst	1,7	4,9	6,1
Änderungen der Ausgaben [%] für ...			
Frischobst	4,7	5,9	6,1
Bananen	13,8	19,9	20,9
Kernobst	3,8	– 0,2	0,2
Anderes Frischobst	1,7	4,9	6,1
Änderungen der Ausgaben [DM (real 1995)/Jahr u. Person] für ...			
Frischobst	6,64	4,61	6,42
Bananen	3,63	3,34	3,81
Kernobst	1,86	– 0,08	0,07
Anderes Frischobst	1,15	1,35	2,54

Quelle: Eigene Berechnungen

3.4.2 Auswirkungen auf Mengen und Ausgaben

Als Basisdaten für die Quantifizierung der Auswirkungen auf die nachgefragten Mengen und die Höhe der Ausgaben wurden die durchschnittlichen Realpreise und Realausgaben der Haushalte im Zeitraum 1994–1998 verwendet (siehe Punkt 3.3.2). Es wurde davon ausgegangen, dass der Bananenpreis durch die Bananenmarktordnung um 38% erhöht (siehe Punkt 3.4.1) wurde, sowie vereinfachend angenommen, dass die Preise der anderen Frischobstkategorien und die gesamten Konsumausgaben nicht beeinflusst wurden. Zuerst wurden mit der ökonometrisch geschätzten Nachfragefunktion für Frischobst (erste Stufe) die nachgefragten Frischobstmengen (als Index) sowie die Ausgaben für Frischobst für die Situationen mit und ohne Bananenmarktordnung bestimmt. Anschließend wurden mit den Regressionsergebnissen der Nachfragesysteme (zweite Stufe) die Ausgabenanteile der verschiedenen Frischobstkategorien berechnet und daraus die Ausgaben und nachgefragten Mengen ermittelt (siehe Tabelle 3.2).

Wegen der unelastischen Frischobstnachfrage blieben die nachgefragten Mengen in etwa konstant, während die Ausgaben für Frischobst durch die Bananenmarktordnung zwischen 4,7% und 6,1% anstiegen. Dagegen wurde bei der Bananennachfrage sowohl ein starker Rückgang der Menge (12,5%–17,6%) als auch eine deutliche Erhöhung der Ausgaben ausgelöst

(13,8%–20,9%), wobei der „ärmere“ Haushaltstyp 1 den Bananenkonsum am stärksten einschränkte. Aufgrund von Substitutionsbeziehungen stieg die Nachfrage nach Kernobst und der Kategorie „anderes Frischobst“ an.

3.5 Schlussbetrachtung

Bei den Verhandlungen zur Gestaltung und Veränderung der EU Bananenmarktordnung spielten die Verbraucherinteressen stets eine untergeordnete Rolle. Zwar entspricht die beschriebene Erhöhung der Frischobstausgaben nur zwischen 0,05% (Haushaltstyp 1) und 0,036% (Typ 3) der Gesamtausgaben der Haushalte, doch da insbesondere die „ärmeren“ Haushalte betroffen sind, ist unter dem Gesichtspunkt der sozialen Gerechtigkeit und des Interessensausgleichs aller Betroffenen zu wünschen, dass bei der 2006 geplanten Umgestaltung der Bananenmarktordnung auch die Verbraucherinteressen berücksichtigt werden. Hinzu kommt, dass nur 22% des Wohlfahrtsverlustes der EU-Verbraucher bei den bevorzugten Produzenten in der EU und in den AKP-Staaten ankommen (Kersten, 1995).

Literaturverzeichnis

Anderson, G. J. und R. W. Blundell (1983): Testing Restrictions in a Flexible Dynamic Demand System: An Application to Consumers' Expenditure in Canada. In: Review of Economic Studies, 50:397–410.

Burrell, A. und A. Henningsen (2001): An Empirical Investigation of the Demand for Bananas in Germany. In: Agrarwirtschaft, 50(4):242–249.

Deaton, A. und J. Muellbauer (1980): Economics and Consumer Behaviour. Cambridge University Press, Cambridge.

Fan, S., E. J. Wailes und G. L. Cramer (1995): Household Demand in Rural China: A Two-Stage LES-AIDS Model. In: American Journal of Agricultural Economics, 77:54–62.

Hallam, D. und L. Peston (1997): The Political Economy of Europe's Banana Trade. Occasional Paper 5, Department of Agricultural and Food Economics, University of Reading, Reading.

Kersten, L. (1995): Impacts of the EU Banana Market Regulation on International Competition, Trade and Welfare. In: European Review of Agricultural Economics, 22:321–335.

Kersten, L. (2000): Zur Revision der Bananenmarktordnung. Unveröffentlicht. Zusammenfassung veröffentlicht im Jahresbericht 1999 der FAL Braunschweig-Völkenrode, Kapitel 5.1, S. 136-137.

Kapitel 4

Abschnitt „Methodisches Vorgehen“ aus „Auswirkungen der Mid-Term-Review-Beschlüsse auf den Agrarsektor und das Agribusiness in Schleswig-Holstein und Mecklenburg-Vorpommern“

Christian H.C.A. Henning, Arne Henningsen, Carsten Struve und Jörg Müller-Scheeßel

erschienen in: Agrarwirtschaft – Sonderheft 178, Agrimedia GmbH, Bergen/Dumme, 2004, S. 95–117.

4.1 Ein regionalisiertes LP-Modell

Es wurde ein einheitliches lineares Programmierungsmodell (LP) erstellt, das die landwirtschaftlichen Betriebe in Schleswig-Holstein bzw. Mecklenburg-Vorpommern abbildet. In einem Linearen Programmierungsmodell wird von verschiedenen Aktivitäten ausgegangen, die bestimmte Deckungsbeiträge generieren, aber gleichzeitig bestimmten Restriktionen unterliegen. Unter Berücksichtigung der einzelnen Deckungsbeiträge und der Restriktionen wird der gesamte Deckungsbeitrag maximiert.

In diesem Modell können die Betriebe zwischen ca. 1100 verschiedenen (Produktions-) Aktivitäten wählen, die in den nächsten Abschnitten näher erläutert werden. Der Umfang der Aktivitäten wird durch die vorhandenen Ressourcen (z.B. Boden, Stallplätze) begrenzt. Eine Optimierung der Betriebsorganisation kann unter dem Basisszenario, d.h. die Situation in den Jahren 2001–2003 („Base-Run“), dem Brüsseler Grundszenario der Agrarreform (Betriebsprämie) und auch unter diversen nationalen Optionen (z.B. Regionalprämie, Teilkopplungen usw.) erfolgen.

Der landwirtschaftliche Sektor Schleswig-Holsteins wird durch 416 Modellbetriebe abgebildet. Diese stehen jeweils für eine Gruppe von Betrieben und unterscheiden sich im Betriebstyp, in der Betriebsgröße und im Naturraum, in dem sie liegen.

Dem Modell für Schleswig-Holstein entsprechend wird auch der landwirtschaftliche Sektor Mecklenburg-Vorpommerns durch 156 Modellbetriebe abgebildet. Diese Betriebe liegen in 12 verschiedenen Landkreisen und unterscheiden sich in der Produktionsausrichtung. Aufgrund unzureichender Statistiken wird hier nicht auf Größenklassen eingegangen.

4.1.1 Modellbetriebe für Schleswig-Holstein

Die Einteilung des Schleswig-Holsteinischen Agrarsektors in Modellbetriebe beruht auf Daten von knapp 15.000 realen landwirtschaftlichen Betrieben, d.h. nahezu alle landwirtschaftlichen Betriebe in Schleswig-Holstein werden berücksichtigt. Für jeden dieser Betriebe sind uns der Unternaturraum, die landwirtschaftliche Nutzfläche, deren Aufteilung in ackerfähiges Land und in Dauergrünland, die prämienberechtigte Fläche (für Grandes-Cultures-Prämien), die Zuckerrübenfläche sowie die Summe der Kartoffel- und Gemüsefläche bekannt. Weiterhin haben wir Informationen über die Milchquote und die Höhe der erhaltenen Tierprämien, aufgeteilt in Mutterkuhprämie, Summe aus Schlacht- und Ergänzungsprämie, Rindersonderprämie für Bullen, Rindersonderprämie für Ochsen und Mutterschafprämie.

Da es auch mit heutigen Computern nicht in angemessener Zeit möglich ist, das Modell für alle 15.000 Betriebe durchzurechnen, haben wir die Betriebe in verschiedene Gruppen eingeteilt:

- 22 Unternaturräume[1],
- 5 Betriebstypen (Marktfrucht, Futterbau Milch, Futterbau Mast, Marktfrucht Futterbau, Futterbau Marktfrucht) und
- 4 Größenklassen (5–60 ha, 60–100 ha, 100–200 ha, $\geq$200 ha)

Die Einteilung der Betriebe in die verschiedenen Betriebstypen wurde anhand der von uns berechneten Standarddeckungsbeiträge der einzelnen Betriebszweige durchgeführt. Betriebe mit weniger als 5 ha wurden nicht betrachtet, da ihre Gesamtfläche nur einen geringen Anteil ausmacht. Es handelt sich bei diesen Betrieben insbesondere um Hobby-Betriebe oder Erzeuger von Spezial- oder Nischenprodukten, so dass ihre Produktionsentscheidungen kaum durch die Agrarreform beeinflusst werden. Aus den 440 betrachteten Betriebsgruppen (22 Unternaturräume $\times$ 5 Betriebstypen $\times$ 4 Größenklassen) haben wir 290 Gruppen ausgewählt, deren landwirtschaftliche Nutzfläche mindestens einen Anteil von 2% am Unternaturraum oder 0,05% an Schleswig-Holstein hat. Von jeder dieser Gruppen haben wir den Durchschnitt gebildet, so dass jede Gruppe durch einen Durchschnitts- bzw. Modellbetrieb repräsentiert wird. Die sektoralen Ergebnisse der einzelnen Naturräume bzw. Schleswig-Holsteins insgesamt können somit ermittelt werden, indem die einzelbetrieblichen Modellergebnisse mit der Anzahl der Betriebe in der entsprechenden Gruppe multipliziert werden.

Aus der durchschnittlichen Milchleistung in den jeweiligen Unternaturräumen wurde die Anzahl der gehaltenen Milchkühe berechnet. Die Anzahl der gehaltenen Mutterkühe bzw. Mutterschafe konnte direkt aus der Mutterkuh- und Mutterschafprämie abgeleitet werden. Dagegen konnten wir aus den Schlacht-, Ergänzungs- und Rindersonderprämien nur die Anzahl der geschlachteten Tiere ableiten. Allerdings war es uns möglich, mit zusätzlichen produktionstechnischen Daten (z.B. Haltungsdauer, Remontierungsrate) auch die Anzahl der übrigen gehaltenen Rinder abzuschätzen. Aus der so berechneten Anzahl der gehaltenen Tiere konnten wiederum die vorhandenen Stallplätze abgeleitet werden. Dabei wurde angenommen, dass bei den Betrieben die Stallkapazitäten zu 97,5% ausgenutzt werden.

Die einzelbetrieblichen Daten enthalten leider keine Angaben über die Schweinehaltung in den Betrieben. Die Schweinehaltung ist zwar nicht direkt von der Agrarreform betroffen, doch übt sie durch die Gülleerzeugung und den Bedarf an Arbeitskräften einen Einfluss auf andere

[1]Schleswig-Holstein ist in 23 Unternaturräume eingeteilt, es werden aber nur 22 Unternaturräume betrachtet, da Helgoland keine landwirtschaftliche Bedeutung hat.

Betriebsbereiche aus. Da es sich hierbei jedoch nur um einen indirekten Einfluss handelt, scheint eine grobe Abschätzung gerechtfertigt. Es wird vereinfachend davon ausgegangen, dass alle Schweine im Betriebstyp „Marktfrucht“ gehalten werden. Aus diesem Grunde wurde eine weitere Unterteilung dieses Betriebstyps in vier zusätzliche Betriebstypen vorgenommen. Alle Größenklassen des Betriebstyps „Marktfrucht“ wurden unternaturraumspezifisch den Betriebstypen „Marktfrucht – extensiv“, „Marktfrucht Veredlung“, „Veredlung Marktfrucht“ und „Veredlung Schwein“ zugeordnet. Diese Aufteilung erfolgte entsprechend den prozentualen Anteilen dieser Betriebstypen, wie sie in den „Wirtschaftsergebnissen 2001/2002“ des Landwirtschaftlichen Buchführungsverbandes veröffentlicht wurden. Aus der dort angegebenen Anzahl an Sauen und Mastschweinen können die jeweiligen Stallplätze ermittelt werden. Da die Summe der so ermittelten Stallplätze auf Unternaturraumebene in vielen Fällen von den vom Statistischen Landesamt ermittelten Tierzahlen abweicht, wird die Anzahl der Stallplätze auf Unternaturraumebene gleichmäßig prozentual erhöht oder verringert, so dass die Tierzahl auf Unternaturraumebene nach der Anpassung genau den statistischen Zahlen entspricht. Diese Aufteilung des Betriebstyps „Marktfrucht“ führte zu einer Erhöhung der Anzahl der Modellbetriebe auf 416.

Das Grünland wurde in Mineralbodengrünland und Niedermoorgrünland aufgeteilt. Die jeweiligen Anteile sind unternaturraumspezifisch und basieren auf Experteneinschätzungen auf Grundlage der gesamten Moorfläche in den jeweiligen Unternaturräumen. Der gesamte Mineralboden (Acker + Mineralbodengrünland) wurde in 10 Bodengüteklassen (20–22; 23–25; 26–29; 30–34; 35–39; 40–44; 45–49; 50–59; 60–69 und ≥70 Bodenpunkte) aufgeteilt. Zunächst wurde für jeden einzelnen Unternaturraum die Verteilung der Güteklassen für den Mineralboden ermittelt. Die Verteilung wurde auf Grundlage der durchschnittlichen Bodenpunkte und der Ackerfläche in jeder Gemeinde des Unternaturraums errechnet. Anschließend wurde diese Verteilung für jeden Modellbetrieb im jeweiligen Unternaturraum zugrunde gelegt. Das Grünland auf Mineralboden wurde zu 75% proportional auf alle Mineralbodengüteklassen verteilt. Die restlichen 25% konnten vom LP-Modell frei auf die Mineralbodengüteklassen verteilt werden. Das Niedermoorgrünland wurde auf 5 Güteklassen in dem Verhältnis 10%, 15%, 25%, 25% und 25% (in aufsteigender Qualität) aufgeteilt.

Für die Aufteilung der gesamten Kartoffel- und Gemüseanbaufläche der Betriebe auf diese beiden Produktionsrichtungen sind leider keine Daten verfügbar. Den überwiegenden Anteil des Gemüseanbaus in den landwirtschaftlichen Betrieben Schleswig-Holsteins stellt der Kohlanbau in Dithmarschen dar. Da sich der Anbau von Kartoffeln und Kohl insbesondere in den Boden- und Fruchtfolgeansprüchen stark unterscheidet, haben wir die Aufteilung in diese beiden Kulturen abgeschätzt. Dazu haben wir für jeden Unternaturraum die von den Betrieben angegebene Kartoffel- und Gemüsefläche mit der vom Statistischen Landesamt veröffentlichten Anbaufläche von Kartoffeln verglichen. Die Kartoffel- und Gemüsefläche wurde so auf die beiden Kulturen aufgeteilt, dass die aufsummierte Kartoffelfläche der Betriebe der gesamten Kartoffelfläche in dem Unternaturraum entspricht. Die restliche Fläche steht den Betrieben als Gemüseanbaufläche zur Verfügung. Im Modell wird vereinfachend davon ausgegangen, dass als einziges Gemüse Kohl angebaut wird.

Da wir über die Ausstattung der Betriebe mit Familienarbeitskräften keine Informationen haben, mussten wir auch hier eine Schätzung vornehmen (siehe Tabelle 4.1). Dabei nehmen wir an, dass im Betriebstyp „Futterbau Milch“ 1,5 Familienarbeitskräfte beschäftigt sind, z.B. der Betriebsleiter vollständig und ein Altenteiler oder sein Ehepartner halbtags. Allen anderen Betrieben haben wir je eine Familienarbeitskraft zugewiesen. Wir gehen davon aus, dass eine Familienarbeitskraft 2000 Stunden pro Jahr effektiv arbeitet. Allerdings kann der Betriebsleiter eines reinen Marktfruchtbetriebes nur 1500 Stunden im Betrieb arbeiten, da er seine Arbeitskraft im Winter nicht auslasten kann.

Tabelle 4.1: Familien-AKh nach Betriebstypen

Betriebstyp	Familien-AK	Familien-AKh
Futterbau Milch	1,5	3000
Futterbau Mast	1	2000
Futterbau Marktfrucht	1	2000
Marktfrucht Futterbau	1	2000
Marktfrucht extensiv	1	1500
Marktfrucht Veredlung	1	2000
Veredlung Marktfrucht	1	2000
Veredlung Schwein	1	2000

4.1.2 Modellbetriebe für Mecklenburg-Vorpommern

Da für Mecklenburg-Vorpommern keine Statistiken über Naturräume vorliegen, wurden hier die 12 Landkreise als geographische Einheiten gewählt.

Die Ermittlung der Faktorausstattung der jeweiligen Betriebstypen auf Ebene der Landkreise führte in Mecklenburg-Vorpommern zu zusätzlichen Problemen, da derartige Daten in keiner Statistik erfasst werden[2]. Es stehen zur näherungsweisen Ermittlung der Faktorausstattung allerdings die beiden folgenden Datensätze zur Verfügung:

[2] mdl. Auskunft Hrubesch, R. und Koblin, G., Stat. Landesamt Mecklenburg-Vorpommern

1. Anzahl der verschiedenen Betriebstypen auf Kreisebene,

2. Anzahl und Faktorausstattung der verschiedenen Betriebstypen auf Landesebene.

Aus diesen beiden Datensätzen kann über die Berechnung der Durchschnittsgröße auf Landesebene und der Anzahl der Betriebe auf Kreisebene die Faktorausstattung auf Kreisebene abgeleitet werden. Diese Methode hat jedoch den Nachteil, dass kreisspezifische Besonderheiten nur dann berücksichtigt werden können, wenn von einem bestimmten Betriebstyp viele Betriebe in einem Landkreis vorhanden sind. Wenn die Faktorausstattung dieser Betriebe deutlich vom Landesdurchschnitt abweicht, kann ein gewisser Fehler nicht ausgeschlossen werden.

Die Faktorausstattung mit Mutterkuh- und Bullenmastplätzen kann auf diese Weise nicht ermittelt werden, da auch hierfür keine statistischen Daten vorliegen. Zur Ermittlung dieser Werte haben wir den Ingenieursansatz herangezogen, in dem die Anzahl dieser Plätze über die freie Futterfläche den Betrieben zugeteilt wird.

Weiterhin mussten die Faktorausstattungen einiger Betriebstypen noch „per Hand“ an die speziellen Bedingungen in den Kreisen modifiziert werden, um z.B. die Ausstattung mit Dauergrünland pro Mutterkuh dem jeweiligen Ertragsniveau anzupassen.

Um weitere betriebsspezifische Charakteristika besser abbilden zu können, wurden einige Betriebstypen noch weiter aufgegliedert (z.B. Milchviehbetriebe mit geringer oder hoher Ausstattung mit Dauergrünland). Somit ergeben sich insgesamt 13 Betriebstypen.

Die Zuteilung der Familienarbeitskräfte ist in Tabelle 4.2 dargestellt und erfolgt analog zu den Betriebstypen in Schleswig-Holstein.

4.1.3 Aktivitäten und ihre Faktoransprüche

Den Betrieben stehen ca. 1100 Aktivitäten zur Verfügung, die durch ca. 550 Restriktionen beschränkt sind. Die produktionstechnischen Daten für diese Aktivitäten basieren auf Planungsdaten[3], betriebswirtschaftlichen Auswertungen[4] und auf Experteneinschätzungen. Ein Vergleich der Ergebnisse des Basisszenarios mit tatsächlich beobachteten Daten diente zur Kalibrierung und Überprüfung des Modells. Je nach Datenverfügbarkeit konnte diese Kalibrierung für Einzelbetriebe, Unternaturräume bzw. Landkreise, Obernaturräume oder den gesamten Agrarsektor des jeweiligen Bundeslandes erfolgen. Im Folgenden wird ein kurzer

[3] z.B. Landwirtschaftskammer Schleswig-Holstein (2002b, 2003a); Landesanstalt für Landwirtschaft des Landes Brandenburg (2001); Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL) (2002)
[4] z.B. Landwirtschaftskammer Schleswig-Holstein (2002a, 2003b,c)

Tabelle 4.2: Familien-AKh nach Betriebstypen

Betriebstyp	Familien-AK	Familien-AKh
Marktfruchtbau extensiv	1,5	1500
Marktfruchtbau mit Schweinen	1	2000
Marktfruchtbau mit Mutterkühen	1	2000
Marktfruchtbau mit Bullenmast	1	2000
Marktfruchtbau mit Milchproduktion (geringer Anteil an Dauergrünland)	1,5	3000
Marktfruchtbau mit Milchproduktion (hoher Anteil an Dauergrünland)	1,5	3000
Futterbau mit Mutterkühen	1	2000
Futterbau mit Bullenmast	1	2000
Futterbau mit Milchproduktion (geringer Anteil an Dauergrünland)	1,5	3000
Futterbau mit Milchproduktion (hoher Anteil an Dauergrünland)	1,5	3000
Veredlung	1	2000
Gemischtbetriebe	1,25	2500
Schafhalter	1	2000

Überblick über die Aktivitäten gegeben. Da eine detaillierte Beschreibung sämtlicher Aktivitäten den Rahmen dieses Abschnitts sprengen würde, wird nur auf die Besonderheiten unseres Modells eingegangen.

4.1.3.1 Pflanzenbau

Im pflanzlichen Bereich stehen den Betrieben alle relevanten Marktfrüchte (diverse Getreidearten, Raps, Leguminosen, Zuckerrüben, Kartoffeln, Kohl), Stilllegungsmöglichkeiten (Phacelia, Selbstbegrünung als Rotation, Dauerbrache, Non-Food-Raps) und Futteranbaumöglichkeiten (Silomais, Ackergras, Acker(gras)mähweide sowie Dauergrünland als Weide, Mähweide oder Wiese) zur Verfügung. Diese Produktionszweige unterscheiden sich je nach Bodenqualität, Vorfrucht und Intensität. Weiterhin können die Betriebe ihr Land mulchen, um vor der Reform Flächen zur Beantragung von Bullenprämien zu haben oder um nach der Reform die Cross Compliance zu erfüllen.

Die Düngung der Pflanzenproduktion kann sowohl mit zugekauften mineralischen Düngemitteln als auch mit wirtschaftseigenem organischen Dünger erfolgen. Dabei werden sowohl die Auflagen der Düngemittelverordnung als auch andere produktionstechnische Beschränkungen des Einsatzes von Wirtschaftsdüngern berücksichtigt. Weiterhin wird eine Auswaschung von Stickstoff- und Kali-Düngern insbesondere bei organischer Düngung und auf leichten Böden beachtet.

Wir gehen davon aus, dass die Preise für sämtliche Marktfrüchte mit Ausnahme von Roggen konstant bleiben (siehe Tabelle 4.3), da sich bei den Interventionspreisen für Getreide nur die monatlichen „Reports“ ändern und die anderen Produktmärkte von der Agrarreform nicht direkt betroffen sind. Aufgrund der Einstellung der Roggenintervention gehen wir davon aus, dass der Erzeugerpreis für Roggen sinken wird. Allerdings wird der Roggenpreis indirekt durch die Gersteintervention noch gestützt, so dass wir nur einen geringen Rückgang von 8,60 auf 8,50 €/dt (ohne MwSt.) angenommen haben.

Da Gerste und Triticale häufig in der eigenen Schweinemast veredelt werden, erhalten Landwirte für dieses Getreide, das sie selbst produzieren und auch in der eigenen Schweinehaltung verfüttern, nicht den Marktpreis, sondern den Futterwert, um der guten Verwertung im eigenen Betrieb Rechnung zu tragen. Da zu erwarten ist, dass nach Aufhebung der Roggenintervention der Roggenpreis so niedrig ist, dass auch eine Verfütterung von Roggen in der eigenen Schweinehaltung attraktiv wird, haben die Betriebe auch die Möglichkeit, Roggen in der eigenen Schweineproduktion zu verfüttern. Da der Futterwert von Roggen aufgrund neuerer Untersuchungen zunehmend höher eingeschätzt wird, haben wir dessen Futterwert auch nach der Reform höher als vor der Reform angenommen. Die Verwendung von eigenem Getreide ist bis zu 1,6 dt pro Mastschwein und 6 dt pro Zuchtsau möglich. Dabei ist der Roggenanteil bei den Mastschweinen auf 30% begrenzt. Der Anteil von Roggen und Triticale zusammen kann bei Mastschweinen maximal 60% und in der Sauenhaltung maximal 15% betragen.

Den Preisunterschied zwischen Food-Raps und Raps zur Verwendung als nachwachsender Rohstoff (NR-Raps) haben wir recht hoch angesetzt, um die erhöhten (z.T. subjektiven) Transaktions- und Risikokosten bei NR-Raps zu berücksichtigen. Diese resultieren unter anderem daraus, dass der Landwirt sich schon lange vor der Ernte in einem Anbauvertrag verpflichten muss, eine bestimmte Mindestmenge („repräsentativer Ertrag“) an NR-Raps abzuliefern. Bei niedrigen Durchschnittserträgen ist das Risiko besonders hoch, diesen „repräsentativen Ertrag“ nicht zu erreichen. Deshalb belegen wir den Anbau von NR-Raps mit einem Risikoabschlag, wenn der langjährige Durchschnittsertrag unter 34 dt/ha liegt. Die Höhe des Risikoabschlages berechnet sich aus der Differenz zwischen dem langjährigen Durchschnittsertrag und 34 dt/ha multipliziert mit 2 €/dt.

Da sich die Preise der meisten Marktfrüchte nicht ändern, ist davon auszugehen, dass sich auch die speziellen Intensitäten nicht ändern. Deshalb stellt das LP-Modell für den Marktfruchtbau für jede Bodenqualität nur eine Intensität bereit, die für diese jeweilige Bodenqua-

Tabelle 4.3: Preise von Marktfrüchten

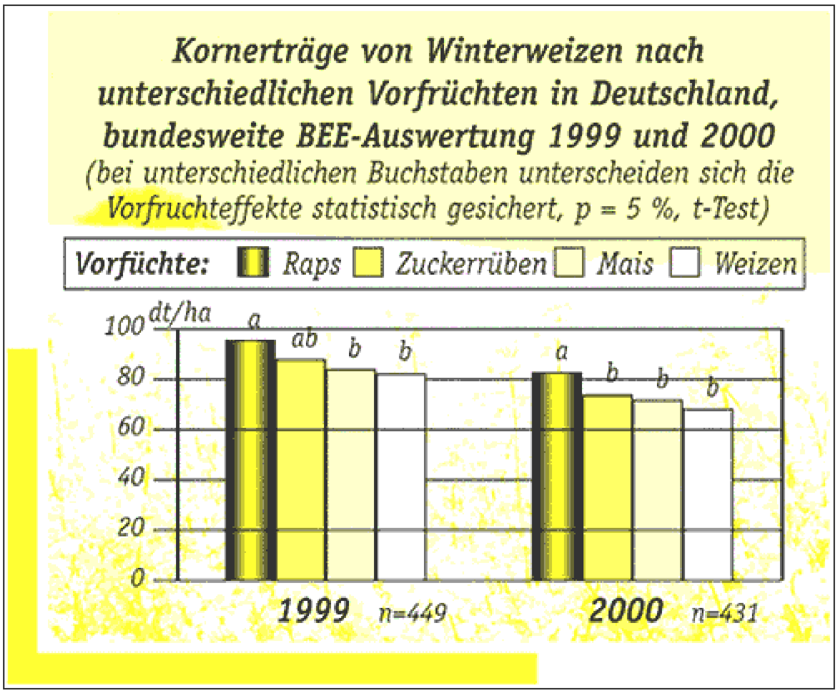
	vor der Reform [€/dt ohne MwSt.]	nach der Reform [€/dt ohne MwSt.]
Weizen	10,00	10,00
Wintergerste	8,60	8,60
Sommerfuttergerste	8,70	8,70
Braugerste	11,50	11,50
Roggen	8,60	8,50
Triticale	8,55	8,55
Hafer	9,00	9,00
Ackerbohnen	11,75	11,75
Raps (incl. 3% Überöl)	21,80	21,80
NR-Raps (incl. 3% Überöl)	20,40	20,40
Zuckerrüben	4,27	4,27
Kartoffeln	8,03	8,03
Kohl	8,00	8,00
Futterwert Wintergerste	8,75	8,75
Futterwert Sommergerste	8,85	8,85
Futterwert Roggen	8,70	8,80
Futterwert Triticale	9,30	9,30

lität optimal ist. Eine Ausnahme hiervon bildet der Roggenanbau, weil der Roggenpreis durch die Aufhebung der Intervention sinken wird. Deshalb stehen im Modell für den Roggenanbau zwei Intensitätsstufen zur Verfügung. Zum einen ist dies der intensive Roggenanbau mit Hybridsorten und zum anderen der extensive Anbau mit Populationssorten.

In unserem LP-Modell werden viele Fruchtfolgerestriktionen explizit berücksichtigt. So hängt z.B. der Weizenertrag stark von der Vorfrucht ab (siehe Abbildung 4.1). Daher haben wir unterschiedliche Weizenanbauverfahren je nach Vorfrucht in das Modell integriert. Wir unterscheiden die Vorfrüchte Raps, Leguminosen, Phacelia, Mais, Stilllegung, Hafer, Ackergras, Zuckerrüben und sonstiges Getreide (insbesondere Weizen).

Beim Rapsanbau erfolgt eine leichte Ertragsminderung, wenn der Rapsanteil in der Fruchtfolge 20% überschreitet, und eine etwas stärkere Ertragsminderung, wenn der Anteil auf über 25% ansteigt (siehe Abbildung 4.2). Der Rapsanteil in der Fruchtfolge kann auf maximal 33% ausgedehnt werden. Darüber hinaus wird im Modell berücksichtigt, dass auch der Zuckerrüben- und Kohlanbau wegen gemeinsamer Fruchtfolgerestriktionen (z.B. Kohlhernie und Rübenzystenälchen) den Rapsanbau einschränkt. Weiterhin wird beachtet, dass der Rapsanbau nach spät räumenden Getreidearten wie z.B. Weizen im Allgemeinen durch eine erhöhte

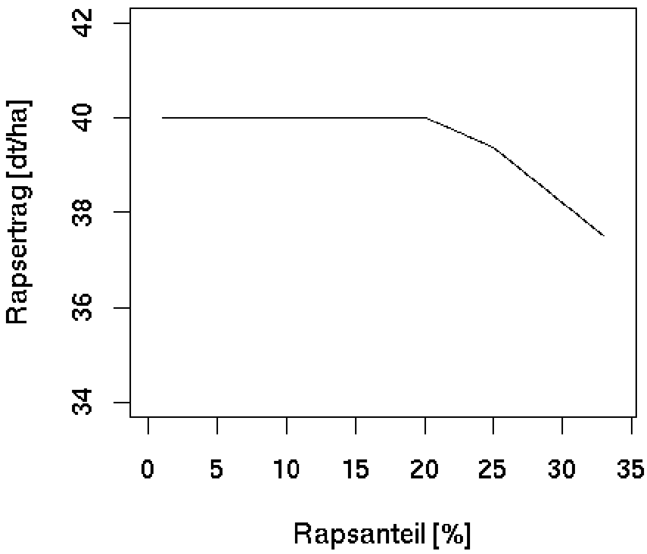
Abbildung 4.1: Weizenertrag nach unterschiedlichen Vorfrüchten



Quelle: UFOP (http://www.ufop.de/1358.htm, Aufruf am 20. 10. 2004)

Saatstärke und eine Herbstdüngung höhere Kosten verursacht und einen leicht niedrigeren Ertrag erreicht als der Rapsanbau nach Wintergerste.

Abbildung 4.2: Modellierte Ertragsdepression bei Raps



Ein wesentlicher Grund zum Anbau von Sommerfrüchten ist, dass es witterungsbedingt teilweise nicht möglich ist, alle Flächen im Herbst zu bestellen. Auch diese Anbaurestriktion haben wir im Modell implementiert. Dabei hängt der Flächenanteil, den der Betrieb im Frühjahr bestellen oder stilllegen muss, von der Vorfrucht und dem Bodentyp ab. Er variiert zwischen 1,3% nach Wintergerste auf Sandboden und 50% nach Zuckerrüben auf schwerer Marsch.

Der Maisanbau kann zum einem als Monokultur und zum anderen innerhalb der Fruchtfolge erfolgen. Dabei wird angenommen, dass der Anbau als Monokultur einen höheren Herbizidaufwand erfordert, wobei nach dem Anbau innerhalb der Fruchtfolge als Folgefrucht entweder ein spät gesäter Weizen, eine Sommerfrucht oder Rotationsbrache (Stilllegung) folgen muss.

Da wichtige Einflussfaktoren beim Anbau von Kartoffeln und Kohl wie z.B. vorhandene Abnahmeverträge und Lagerkapazitäten nur schwer modelliert werden können, haben wir die maximale Fläche für den Kartoffel- bzw. Kohlanbau exogen vorgegeben. Diese maximale Fläche entspricht jeweils genau der Fläche, welche die Modellbetriebe tatsächlich für den Kartoffel- und Kohlanbau heute nutzen. Da der Deckungsbeitrag für diese Früchte wegen hoher versunkener Kosten (z.B. Bau einer Lagerhalle) recht hoch ist, wird diese maximale Fläche immer ausgenutzt.

Da nach der Agrarreform auch Zuckerrübenflächen prämienberechtigt sind, wird der Anbau von „C“-Zuckerrüben wirtschaftlicher. Dies könnte zu einer Ausdehnung der Rübenanbaufläche führen. Andererseits ist zu erwarten, dass im Rahmen der EU-Zuckermarktreform die Rübenpreise sinken und/oder die Rübenquoten gekürzt werden. Dies hätte den gegenläufigen Effekt einer Verringerung der Rübenanbaufläche zur Folge. Da bisher noch offen ist, in welcher Form die EU-Zuckermarktordnung reformiert werden wird, gehen wir vereinfachend davon aus, dass die Rübenanbaufläche der Betriebe konstant bleibt.

Der Anbau von Ackergras kann nur auf Ackerland erfolgen, wobei die Grasnarbe alle vier Jahre umgebrochen wird und eine Neuansaat auf derselben oder einer anderen Fläche erfolgt. Dies führt zu höheren Kosten, aber auch zu einem deutlich höheren Ertrag als der Grasanbau auf Dauergrünland. Neben dem absoluten Dauergrünland kann auch das Ackerland als Weide, Mähweide oder Wiese genutzt werden. Für die Weidenutzung stehen 4 Intensitätsstufen zur Verfügung. Die Mähweiden können ein-, zwei- oder dreimal geschnitten werden, wobei sie jeweils anschließend beweidet werden. Auch die Wiesen können unterschiedlich intensiv genutzt werden, indem sie ein-, zwei-, drei- oder viermal geschnitten werden.

In unserem Modell werden für die einzelnen Produktionsaktivitäten die Saatgut-, Düngemittel-, Pflanzenschutz-, Maschinen-, Trocknungs- und sonstigen Kosten berücksichtigt. Weiterhin werden Zinsansätze für das Umlaufkapital als kalkulatorische Kosten berechnet.

Für die Maschinenkosten berechnen wir die Vollkosten, also sowohl die variablen Kosten als auch die Abschreibung der Maschinen. Die landwirtschaftliche Nutzfläche wird in unserem Modell exogen vorgegeben. Die Betriebe können also kein weiteres Land zu- bzw. verpachten. Deshalb haben etwaige Pachtzahlungen als Fixkosten keinen Einfluss auf die Produktionsentscheidungen und brauchen folglich in unserem Modell nicht berücksichtigt zu werden.

Der Basislauf hat ergeben, dass ein Teil des Grünlandes nicht genutzt wird. Das kann jeweils unterschiedliche Gründe haben. Zum Teil haben Betriebe nicht unerhebliche Flächen an Dauergrünland, aber keine oder nur sehr wenig Milchquote, Mutterkühe, Bullen oder Schafe. Daher weist unser Modell diesen Betrieben keine oder nur wenige Stallplätze zu, so dass sie nicht genügend Tiere halten können, um sämtliches Grünland zu nutzen. Da es in Schleswig-Holstein aber nur sehr wenig ungenutztes Grünland gibt, ist anzunehmen, dass dieses Grünland anderweitig genutzt wird. Dies kann z.B. Pferdehaltung oder Haltung von „Hobby-Mutterkühen“ ohne Prämien sein. (Es gibt in Schleswig-Holstein ca. 50.000 Pferde und Ponys sowie ca. 20.000 Mutterkühe, für die keine Mutterkuhprämie gezahlt wird und deren Haltung somit wirtschaftlich unrentabel ist.) Das vom Modell im Basislauf als überschüssig angesehene Grünland wird als „Hobby“ deklariert, und die gleiche Fläche muss im Reformlauf wieder als „Hobby“ verwendet werden.

4.1.3.2 Tierproduktion

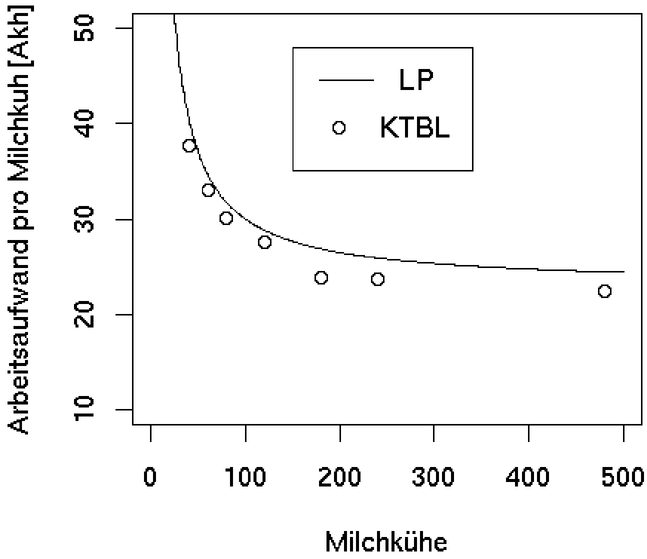
Bei der Tierproduktion können die Betriebe die Produktionsrichtungen Milchviehhaltung, Kälbermast (bis ca. 280 kg Lebendgewicht), Bullen-Intensivmast, Bullen-Weidemast, Mutterkuhhaltung, Sauenhaltung, Schweinemast und Schafhaltung aufnehmen.

Um die Skaleneffekte in der Milchviehhaltung abzubilden, ist der Arbeitseinsatz pro Milchkuh nicht konstant, sondern besteht aus einem Sockelbetrag von 700 AKh/Jahr für die Milchproduktion an sich und zusätzlich 23 AKh/Jahr pro Milchkuh. Diese Spezifikation ist von den Daten des KTBL (Kuratorium für Technik und Bauwesen in der Landwirtschaft)[5] abgeleitet (siehe Abbildung 4.3). Da die Daten des KTBL für sehr moderne und effiziente Betriebe gelten, haben wir für unsere Durchschnittsbetriebe einen etwas höheren Arbeitsbedarf angenommen.

Die Milchleistung in den jeweiligen Unternaturräumen Schleswig-Holsteins wurde aus der „Betriebszweigabrechnung 2001/2002“ des Landwirtschaftlichen Buchführungsverbandes entnommen. Diese wurde noch jeweils um 7% gekürzt, weil in dieser Quelle hauptsächlich die besseren Betriebe ausgewertet wurden (Durchschnitt SH = 6450 kg/Kuh; Durchschnitt LBV = 6948 kg/Kuh). Die Milchleistung in den verschiedenen Landkreisen Mecklenburg-Vorpom-

[5] siehe Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL) (2002, S. 202)

Abbildung 4.3: Modellierter jährlicher Arbeitsaufwand in der Milchviehhaltung



merns wurde auf Grundlage der im Agrarbericht des Landes veröffentlichten Milchertragsklassen ermittelt. Es wird bis 2013 eine Milchleistungssteigerung in Höhe von 10% unterstellt.

Sowohl der Grundfutterbedarf als auch die variablen Kosten der Milchviehhaltung (z.B. Kraftfutter, Tierarztkosten) hängen von der Milchleistung ab. Von der erzeugten Milch werden 98,5% verkauft und setzen daher eine entsprechende Milchquote voraus. Die übrigen 1,5% werden im Haushalt verbraucht oder zur Kälberfütterung verwendet. Auch für diese Milch erhalten die Betriebe den Milchpreis, da in der Färsenaufzucht und Bullenmast schon die Kosten für Milch und Milchaustauscher berücksichtigt sind. Wegen der Milchpreissenkung reduzieren sich im Reformszenario die Kosten für Milch in der Kälberaufzucht um 10 €/Tier.

Der Grundfutterbedarf von allen Wiederkäuern wurde in 5 Teile aufgeteilt: 4 Weideperioden und Winterfutter. Dabei kann das Weidefutter durch Winterfutter ersetzt werden, aber das Weidefutter in einer Periode kann weder das Weidefutter in einer anderen Weideperiode noch das Winterfutter ersetzen. Bei Milchkühen ist auch eine Zufütterung von Winterfutter (insbesondere Silomais) im Sommer vorgesehen. Des Weiteren wird sichergestellt, dass der Strukturfutteranteil im Grundfutter in jeder Periode ausreichend hoch ist.

Die von uns verwendeten Preise der Tierprodukte sind in Tabelle 4.4 dargestellt. Wir gehen davon aus, dass der Grundpreis der Milch von 29,1 auf 21,9 €Ct/kg (ohne MwSt. und ohne Zuschläge für Fett und Eiweiß) fallen wird. Es wird ein Zuschlag für Fett und Eiweiß in Höhe von 5% des Grundpreises gewährt. Diesen Zuschlag haben wir prozentual gewählt, da wir vermuten, dass er sich in etwa proportional zum Grundpreis verringern wird.

Weiterhin nehmen wir an, dass die Preise für Rind-, Lamm- und Schweinefleisch sowie für Ferkel in etwa konstant bleiben. Dagegen werden die Preise für den Lebendverkauf von Nutzrindern aufgrund der Entkopplung der Schlacht-, Ergänzungs- und Rindersonderprämien sinken. Die Preise von Bullkälbern wurden so gewählt, dass in Schleswig-Holstein die Anzahl der verkauften Bullkälber in etwa der Zahl der gekauften Bullkälber entspricht. Dieser Gleichgewichtspreis wurde sowohl vor der Reform als auch nach der Reform ermittelt, so dass die Auswirkungen auf den Kälberpreis eingeschätzt werden können.

4.1.3.3 Prämien

Weiterhin stehen den Betrieben alle relevanten Prämienzahlungen zur Verfügung. Die spezifischen Prämienregelungen wurden sowohl vor als auch nach der Reform exakt abgebildet. Da die Bodenqualität in einem Unternaturraum wesentlich stärker variiert als in einem einzelnen landwirtschaftlichen Betrieb, würden die Modellbetriebe in einem Standard-LP-Modell nur die Böden mit der schlechtesten Qualität in einem Unternaturraum stilllegen. Dies ist realen Einzelbetrieben nicht möglich. Daher wurde den Modellbetrieben vorgegeben, dass jeweils mindestens 80% der Stilllegungsfläche die gleiche Bodenqualität aufweisen muss wie die Fläche der Anbauaktivitäten, die diese Stilllegung erfordern.

Da viele Futterbaubetriebe nur eine geringe Grandes-Cultures-Fläche haben, wurde auch die „Kleinerzeugerregelung“ in das LP-Modell integriert. Im Basislauf entscheidet das LP-Modell, ob die Kleinerzeugerregelung in Anspruch genommen wird, je nachdem, ob sie zu einem höheren oder niedrigeren Gesamtbetriebsdeckungsbeitrag führt. Nach der Agrarreform haben die Betriebe keinen Entscheidungsspielraum mehr. Beim Brüsseler Betriebsmodell sind genau die Betriebe Kleinerzeuger, die auch in der Basisperiode Kleinerzeuger waren. Bei einer Regionalisierung sind die Betriebe als Kleinerzeuger modelliert worden, die in Schleswig-Holstein mehr als 13,51 ha und Mecklenburg-Vorpommern mehr als 15,1 ha Ackerland haben[6].

Im Basislauf erfolgt eine 1%ige „Modulation“ der Prämienzahlungen oberhalb von 10.000 €/Betrieb. Dies entspricht dem Mittelwert aus den Jahren 2002 und 2003 mit 0% bzw. 2% Modulation. Im Reformszenario werden die Prämien oberhalb von 5.000 €/Betrieb um 5% moduliert.

4.1.3.4 Sonstige Aktivitäten

Die Betriebe können sowohl Arbeit zukaufen als auch verkaufen. Dabei werden den Betrieben Kosten in Höhe von 10 € für jede zugekaufte Arbeitskraftstunde (AKh) berechnet. Diese

[6] Es wurden die alten Kleinerzeugergrenzen verwendet, da den Autoren zum Zeitpunkt der Modellerstellung nicht die jetzt angewendeten Flächengrenzen bekannt waren.

Tabelle 4.4: Preise in der Tierhaltung

	Schlacht-gewicht [kg]	Preis pro kg [€/kg incl. MwSt.]		Preis pro Tier [€ incl. MwSt.]	
		vor der Reform	nach der Reform	vor der Reform	nach der Reform
Milch		0,33	0,25		
Altkuh Schwarzbunt	300	1,40	1,40	420	420
Altkuh Fleischrind	348	1,50	1,50	522	522
Bulle Schwarzbunt (Intensivmast)	341	2,10	2,10	715	715
Bulle Schwarzbunt (Weidemast)	324	2,00	2,00	648	648
Färse Schwarzbunt	270	2,00	2,00	540	540
Bulle Fleischrind	412	2,25	2,25	928	928
Färse Fleischrind	280	2,20	2,20	616	616
Mastfärse Fleischrind	343	2,15	2,15	738	738
Mastkalb	147	2,20	2,20	324	324
Absetzer Fleischrind weibl. schlachten	144	2,20	2,20	317	317
Absetzer Fleischrind weibl. verkaufen	*240	*1,40	*1,26	336	302
Absetzer Fleischrind männl. schlachten	170	2,30	2,30	391	391
Absetzer Fleischrind männl. verkaufen	*270	*2,25	*1,69	608	456
Kuhkalb Schwarzbunt verkaufen				30	1
Kuhkalb Schwarzbunt zukaufen				60	20
Bullkalb Schwarzbunt verkaufen				102	16
Bullkalb Schwarzbunt zukaufen				127	36
Lamm	*45	*1,80	*1,80	81	81
Mastschwein	95	1,35	1,35	128	128
Ferkel	28			52	52

* = Lebendgewicht bzw. Preis pro kg Lebendgewicht

Tabelle 4.5: Betriebs- und Regionalprämien

	Betriebsprämie		Regionalprämie		
	Prämien-recht	Prämien-akti-vierung	Prämien-recht Acker	Prämien-recht Grünland	Prämien-akti-vierung
Getreide, Raps	X	X	X		X
Zuckerrüben		X	X		X
Kartoffeln, Kohl			X		X*
Silomais	X	X	X		X
Ackergras (Futterfläche**)	X	X	X		X
Dauergrünland (Futterfläche**)	X	X		X	X
Dauergrünland (keine Futterfläche**)		X		X	X

* maximal bis zur historischen Fläche
** Futterfläche für Rinder oder Schafe, die im Grundantrag ausgezeichnet ist

Arbeit kann durch fest angestellte Arbeitnehmer oder durch nicht-ständige Arbeitskräfte wie z.B. Erntehelfer ausgeübt werden. Es ist aber auch möglich, dass diese Arbeit von Lohnunternehmen bereitgestellt wird. Die Kosten für den Lohnunternehmer setzen sich aus den Arbeitslöhnen und den Maschinenkosten zusammen. Letztere sind dabei schon als Vollkosten in den Produktionskosten enthalten. Wir nehmen also vereinfachend an, dass es genauso teuer ist, Maschinen selbst zu unterhalten und einen Angestellten die Arbeiten verrichten zu lassen wie einen Lohnunternehmer mit der Arbeit zu beauftragen.

Die Opportunitätskosten für die Arbeit von Familienarbeitskräften variieren sehr stark. Während sie für junge und gut ausgebildete Arbeitskräfte hoch sein können, sind sie insbesondere für ältere Arbeitskräfte, die in der heutigen wirtschaftlichen Lage wenig Chancen auf dem Arbeitsmarkt haben, annähernd Null. Weiterhin muss man von dem möglichen Arbeitslohn außerhalb der Landwirtschaft einen Abschlag für die Präferenz von landwirtschaftlicher Arbeit vornehmen. Aufgrund dieser Überlegungen halten wir durchschnittliche Opportunitätskosten in Höhe von 3 €/AKh für realistisch und benutzen diesen Wert daher auch in unserem Modell. Langfristig (>15 Jahre) liegt dieser Wert im Rahmen des Generationenwechsels natürlich höher.

In unserem Modell werden Pachtzahlungen sowie Kreditzinsen nicht berücksichtigt, so dass man von den von uns ermittelten Gesamtbetriebsdeckungsbeiträgen diese Zahlungen noch abziehen muss, um auf den Betriebsgewinn zu kommen. Da keine umfassenden Informationen über die Pachtanteile und die Fremdkapitalbelastung der Betriebe vorliegen, ist nur eine Abschätzung der Betriebsgewinne möglich.

4.1.4 Gewinnermittlung

Ausgehend vom Gesamtdeckungsbeitrag, der im LP errechnet wurde, wird eine Näherung für den Gewinn errechnet.

Vom Gesamtdeckungsbeitrag, der auch nichtlandwirtschaftliche Arbeitseinkünfte enthält, werden die Aufwendungen für die Unterhaltung und Abschreibungen der Gebäude und Grundverbesserungen abgezogen, weil diese nicht in den Deckungsbeitragsrechnungen des LP berücksichtigt werden. Da diese Daten nicht betriebsindividuell bekannt sind, werden die Daten des Landwirtschaftlichen Buchführungsverbandes Schleswig-Holstein als Näherungswert angenommen. Dabei werden betriebstypenspezifische Werte gewählt, so dass unterschiedliche Fixkosten entstehen. Des Weiteren werden die Pachtkosten abgezogen.

Zur Ermittlung der Pachtkosten wird folgendes unterstellt:

Die Pachtanteile von Grünland und Ackerland auf Unternaturraumebene sollen den Werten in der Statistik entsprechen. Die Pachtanteile in größeren Betrieben werden höher als bei kleineren Betrieben angenommen.

Zunächst wird von den in Tabelle 4.6 genannten Pachtanteilen ausgegangen.

Tabelle 4.6: Angenommene Pachtanteile

Größenklasse	angenommener Pachtanteil
1 (bis 60 ha)	35%
2 (60 bis 100 ha)	40%
3 (100 bis 200 ha)	45%
4 (über 200 ha)	50%

Die sich daraus ergebenden regionalen Pachtanteile werden mit den statistischen Daten auf Unternaturraumebene (UN) abgeglichen. Der betrieblich kalibrierte Pachtanteil berechnete sich dann aus:

$$\text{kalibrierter Pachtanteil}_i = \text{angenommener Pachtanteil}_i \times \text{Faktor} \tag{4.1}$$

mit $\text{Faktor} = \dfrac{\text{Pachtanteil}_i \text{ laut Statistik im UN}}{\text{Pachtanteil}_i \text{ im UN bei angenommenen Pachtanteilen der Größenklasse}}$
mit i = Ackerland, Grünland

Es wird davon ausgegangen, dass bei allen Acker- und Grünlandböden der Pachtanteil unabhängig von der Qualität jeweils gleich ist. Des Weiteren wird unterstellt, dass der Pachtpreis 80% des regional durchschnittlichen Schattenpreises entspricht. Dies entspricht einem Abschlag von 20% für Risiko und hektarbezogene Grundkosten wie Berufsgenossenschaft etc.

Die Pachtzahlungen werden berechnet, indem die Pachtflächen des Betriebes mit den qualitätsspezifischen Pachtpreisen multipliziert werden.

Ein weiterer Kostenpunkt sind die Pachten für Milchquoten. In Schleswig-Holstein beträgt der Anteil der gepachteten Quote etwa 25% in 2003.

Sofern die Milchprämie an das Land gebunden wird (Szenarien 3, 5 und 7), sinken die Quotenkosten zumindest um den Wert der Milchprämie. Wenn allerdings die Milchprämie an die Quote gebunden bleibt, fiele der Wert des Milchprämienrechts dauerhaft an den Quotenbesitzer, so dass bei Verlängerung von Pachtverträgen oder Quotenkauf auch für die Prämie zu bezahlen wäre (Szenarien 4 und 6).

4.1.5 Implementierung des Modells

Die Programmierung dieses Modells sowie die Auswertung der Modellergebnisse wurde in der freien Statistiksprache und -umgebung „R“ (R Development Core Team, 2005, siehe http://www.r-project.org) implementiert. Dabei erfolgte die Lösung der linearen Programmierungsmodelle mit den R-Paketen „lpSolve“ (Berkelaar and Buttrey, 2004) und „linprog“ (Henningsen, 2003), die intern die LP-Software „lp_solve“ (Berkelaar et al., 2003) verwenden. Zur Erstellung der Landkarten wurden die R-Pakete „shapefiles“ (Stabler, 2003), „maptools“ (Lewin-Koh and Bivand, 2004), „RColorBrewer“ (Neuwirth, 2004) und „pixmap“ (Bivand et al., 2004) benutzt. Die HTML-Tabellen wurden mit dem R-Paket „R2HTML“ (Lecoutre, 2004) erstellt.

4.1.6 Diskussion des Modellansatzes

Dieses Modell wurde konstruiert, um die kurz- und mittelfristigen Auswirkungen der jüngsten Agrarreform zu analysieren. Die langfristige Sektorentwicklung kann mit diesem Modell allerdings nicht abgeschätzt werden, da der Strukturwandel und langfristige Investitionen in Gebäude nicht berücksichtigt werden. Es ist also kein dynamisches, sondern ein komparativ statisches Modell. Weiterhin wurde im Modell keine Intensivierung oder Extensivierung des Marktfruchtbaus implementiert, da der MTR kaum Auswirkungen auf die Marktordnungspreise von Marktfrüchten hat. Daher überschätzen unsere „Simulationsszenarien“ insbesondere die Auswirkungen von Getreidepreissenkungen.

4.2 Das CAPRI-Modell

Da die Auswirkungen des MTR in dieser Studie nicht nur für die beiden nördlichen Bundesländer, sondern für Deutschland, die EU und die gesamte Welt untersucht werden sollen, ist die Erstellung eines eigenen Modells, das die genannten Regionen abbildet, zu aufwändig. Es wird daher auf ein von der EU finanziertes und von der Universität Bonn erstelltes Modell zurückgegriffen. Dieses Modell ist in der Lage, die EU-weite Landwirtschaft detailliert abzubilden, politische Maßnahmen, wie Direktzahlungen an Betriebe oder Eingriffe an der Grenze einzubeziehen und darüber hinaus weltweite Implikationen der Gemeinsamen Agrarpolitik der EU zu berücksichtigen.

Im Folgenden wird das verwendete Modell kurz erläutert und die wichtigsten Annahmen diskutiert. Die Ausführungen in diesem Kapitel stützen sich hauptsächlich auf die folgenden Quellen: Institute for Agricultural Policy, University Bonn et al. (1999); Britz et al. (2003, 2004).

Das CAPRI-Modellsystem ist ein komparativ-statisches Simulationsmodell für den Agrarsektor und dient hauptsächlich für Ex-post- und mittelfristige Ex-ante-Analysen. Es besteht aus einem Angebots- und einem Marktmodul, die über einen iterativen Prozess miteinander gekoppelt sind und in diesem Kapitel getrennt erklärt werden.

4.2.1 Das Angebotsmodul

In dem Angebotsmodul wird für jede der über 200 NUTS-II-Regionen[7] in der EU ein aggregiertes Programmierungsmodell gerechnet, in dem der Gewinn maximiert wird. Diese NUTS-II-Regionen sind in Abbildung 4.4 für die EU-15 und Norwegen dargestellt.

Das Modell umfasst 60 verschiedene Endprodukte aus 50 Produktionsprozessen und 35 Vorleistungsgütern. Das Angebot wird in einem zweistufigen Prozess gerechnet. Zuerst wird der optimale Einsatzumfang der variablen Inputs, die sog. optimale spezielle Intensität, festgelegt. Hierbei wird der Deckungsbeitrag berechnet. In einem zweiten Schritt wird in den nicht-linearen aggregierten Programmierungsmodellen der gewinnmaximierende Produktionsmix berechnet. Hierbei werden die Preise exogen vorgegeben, die in dem Marktmodul generiert werden, welches im Anschluss besprochen wird. Gleichzeitig mit der Berechnung des optimalen Produktionsmixes werden die Futter- und Düngekosten optimiert. Beschränkungen bezüglich des Grün- und Ackerlandes, Stilllegungsverpflichtungen, Quoten sowie der

[7] Die NUTS-Einteilung ist eine Nomenklatur der EU. Jede NUTS-Ebene entspricht in den Mitgliedsländern ca. gleich großen Verwaltungseinheiten. Mecklenburg-Vorpommern und Schleswig-Holstein stellen je eine NUTS-II-Region dar, in Niedersachsen sind die noch bestehenden Bezirksregierungen je eine NUTS-II-Einheit.

Abbildung 4.4: NUTS-II-Regionen in der EU-15 und Norwegen

Quelle: Universität Bonn (CAPRI-Modell), 2004

Futterbedarf der Tiere und der Düngebedarf der Pflanzen werden als Beschränkungen in das Modell aufgenommen, wobei der Düngebedarf durch organische und mineralische Dünger gedeckt werden kann. Futter wird als nicht handelbar abgebildet, somit wird die Tierproduktion an die regionale Futterproduktion gebunden.

Die Ertragsentwicklungen werden exogen durch eine Trendanalyse auf Basis der Jahre von 1980–1999 vorgegeben.

Die Ergebnisse der Programmierungsmodelle auf NUTS-II-Ebene werden daraufhin auf der Ebene des Mitgliedslandes aggregiert und auf die in dem Basisjahr beobachtete Produktionsmenge kalibriert. Dieses geschieht mittels Positiver Mathematischer Programmierung über die Veränderung einer Kostenfunktion, so dass für den im Basisjahr beobachteten Produktionsumfang die Optimalbedingung Preis = Grenzkosten gilt.

Der Einsatz eines Programmierungsmodells erlaubt die direkte Modellimplementierung von Direktzahlungen unterschiedlicher Ausgestaltung, Stilllegungsverpflichtungen und Quoten (vgl. Jansson et al., 2003), sowie die realitätsnahe Abbildung wichtiger Produkt-Produkt-Beziehungen.

Das in den Programmierungsmodellen errechnete Angebot wird daraufhin an das Marktmodul gegeben, welches markträumende Preise für jedes Mitgliedsland generiert. Diese Preise werden dann wieder an das Angebotsmodul zurückgegeben, welches dann das Angebot mit diesen Preisen erneut berechnet. Dieser iterative Prozess geschieht, bis ein Gleichgewicht erreicht ist.

4.2.2 Das Marktmodul

Das Marktmodul bildet zwölf Regionen[8] der Welt und die Mitgliedsländer der EU mittels Funktionen für das Angebot, die Konsum-, Futter- und Verarbeitungsnachfrage ab. Die Parameter dieser Funktionen werden aus den Elastizitäten anderer Studien und Modelle abgeleitet. Diese Funktionen werden mithilfe extern projizierter Mengen und Preise in dem Simulationsjahr kalibriert. Die projizierten Daten für Nicht-EU-Regionen stammen u.a. aus Quellen der FAO (z.B. Bruinsma, 2003). Die Wahl der Funktionsformen (Normalized Quadratic Funktionen für das Angebot und die Futternachfrage, Generalized Leontief-Funktionen für den Konsum) in Kombination mit der Einhaltung bestimmter Bedingungen (Homogenität der Nachfragefunktionen vom Grade Null in Preisen, Symmetrie, Krümmung) sorgen für die notwendige mikroökonomische Fundierung.

In dem Marktmodell können verschiedene Politiken abgebildet werden. Dieses sind bilaterale Zölle (Wert- und Stückzölle), bilaterale Abkommen, globale und bilaterale Einfuhrquoten, sowie die PSE- und CSE-Maße als Ausdruck verschiedener Politiken. Um den Handel zwischen verschiedenen Ländern abzubilden, bestehen grundsätzlich zwei Möglichkeiten:

1. Der Weltmarkt stellt eine Art „Pool“ dar, Exporteure liefern in diesen Pool, Importeure entnehmen die Waren aus diesem. Handelsströme zwischen verschiedenen Ländern sind bedeutungslos.

2. Die zweite Möglichkeit ist die Anwendung der Armington-Annahme. Diese berücksichtigt die Warenströme zwischen verschiedenen Ländern, indem die Produkte nicht nur nach Art (z.B. Weizen) sonder auch nach Herkunft differenziert werden (z.B. Weizen aus Kanada). Für den Handel der einzelnen Weltregionen wird im CAPRI-Modell die Armington-Annahme auf zwei verschiedenen Stufen genutzt. Auf der oberen Stufe wird bestimmt, aus welchen Anteilen an importierten und heimischen Produkten die Gesamtnachfrage in einem Land befriedigt wird. Auf der unteren Ebene werden die

[8] EU-15, Osteuropäische Beitrittskandidaten, Mittelmeerregion, USA, Kanada, Australien und Neuseeland, Entwicklungsländer mit Freihandel, Hochzollländer, Indien, China, AKP-Staaten, Rest der Welt

Importanteile der einzelnen Länder festgelegt. Hierbei wird durch die Wahl der Substitutionselastizitäten bestimmt, dass die Substitution zwischen den einzelnen Importströmen eines Gutes in einem Land im Vergleich zu anderen allgemeinen quantitativen Modellen relativ elastisch ist. Aufgrund der Armington-Annahme berechnet das Modell keinen einheitlichen Weltmarktpreis.

Die Datenbasis ist ein fester Bestandteil des Modells und besteht teilweise aus der Regio-Datenbank der Eurostat und teilweise aus Daten des FADN (Farm Accountancy Data Network). Daten zur Inputallokation entstammen teilweise noch der SPEL[9]/EU-Datenbank, die seit 1997 nicht mehr aktualisiert wird.

4.2.3 Diskussion des Modellansatzes

Das Modell beinhaltet unterschiedliche Annahmen, die an dieser Stelle kurz diskutiert werden sollen:

Die landwirtschaftliche Produktion ist durch die Abbildung einer NUTS-II-Region durch ein Programmierungsmodell stark aggregiert und kann daher die Wirkung auf einzelne Betriebstypen nicht erfassen. Eine detaillierte Abbildung, die den Unterschied zwischen den einzelnen Naturräumen oder einzelner Betriebstypen (vgl. Balmann et al., 1998) erfasst, würde die Berechnung des Modells deutlich erschweren (vgl. Britz et al., 2004) und vermutlich an der Datenbeschaffung scheitern.

Strukturwandel ist in dem Modell nicht abzubilden, da es ein aggregiertes Modell ist und keine Faktormärkte modelliert werden.

Grundsätzlich wird von einer linearen Beziehung zwischen Output und Input ausgegangen, d.h. eine Outputsteigerung von 10% bedingt eine Inputsteigerung von z.B. Energie um ebenfalls 10%. Für diese Inputs wird eine jährliche Einsparung von 0,2% angenommen.

In dem Marktmodell wird durch die Armington-Annahme bewirkt, dass nur in dem Basisjahr beobachtete Handelströme durch die Modellberechnungen erfasst werden, weil eine Kalibrierung auf eine Handelsmenge von Null nicht möglich ist. Es können demzufolge keine neuen Handelsströme in den Modellberechnungen berücksichtigt werden, wenn sie nicht im Basisjahr schon zu beobachten gewesen sind.

Die hier genannten Probleme sind nicht unerheblich. So kann die starke Aggregierung in Kombination mit dem nicht abzubildenden Strukturwandel zu Ergebnisveränderungen im Vergleich zu disaggregierten Modellen führen. Diese disaggregierten Modelle können wie-

[9]Sektorales Produktions- und Einkommensmodell der Landwirtschaft, Modell an der Universität Bonn mit einer Datenbank seit 1978

derum aus den genannten Gründen der Datenbeschaffung und Berechnungsproblemen die Markteffekte nicht gut abbilden und hier wieder Nachteile aufweisen.

Die Probleme im Bereich des Marktes sind hauptsächlich für die verschiedenen Marktszenarien von Bedeutung und resultieren aus den Anforderungen an das Modell, in dem einzelne Handelsströme abgebildet werden sollen. Hierfür besteht keine andere Möglichkeit als die mehr als 30 Jahre alte Armington-Annahme zu unterstellen. Von daher müssen die Probleme bei der Ergebnisauswertung berücksichtigt werden.

4.3 Modellszenarien im LP

Zur umfassenden Analyse der Auswirkung des MTR wurde das regionalisierte LP-Modell neben einem Base-run-Szenario für insgesamt 11 Szenarien gerechnet. Das Base-run-Szenario dient als Referenzszenario für die einzelnen MTR-Szenarien. Alle MTR-Szenarien wurden für das Jahr 2013, d.h. nach Abschluss der Implementationsphase, modelliert.

4.3.1 Base-Run-Szenario

Das Base-run-Szenario umfasst die ökonomischen und politischen Rahmenbedingungen sowie die betrieblichen Verhältnisse (Produktionstechnologie und Ressourcenausstattung), wie sie sich durchschnittlich in der Referenzperiode 2001–2003 vor dem MTR darstellten. Im Einzelnen wurden die folgenden Annahmen für das Base-run-Szenario getroffen.

4.3.2 Ausgestaltung der Entkopplung

Entsprechend Kapitel 2 ergeben sich hinsichtlich der nationalen Ausgestaltung der Entkopplung grundsätzlich unterschiedliche Möglichkeiten. Einerseits kann nach dem Betriebs- bzw. dem Regionalmodell vorgegangen werden. Für das Regionalmodell kann eine einheitliche Flächenprämie oder aber eine getrennte Grünland- bzw. Ackerprämie eingeführt werden (siehe Kapitel 2 oben). Dies ergibt drei Entkopplungsszenarien. Zusätzlich ist es im Regionalmodell möglich, dass die Milchprämie als individuelle Betriebsprämie ausgezahlt wird. Insofern ergeben sich insgesamt die in Tabelle 4.7 dargestellten fünf bzw. sechs Entkopplungsszenarien.

Im Folgenden werden die einzelnen Szenarien kurz beschrieben:

Szenario 2: Entkopplung nach dem Betriebsmodell, wobei die Milchprämie an die Quote gebunden ist.

Tabelle 4.7: Entkopplungsszenarien

	Betriebsmodell	Regionalmodell (Option B)	
Milchprämie		differenziert nach Acker- und Grünland	einheitlich
an Quote gekoppelt	2,3	6	4
auf Fläche bezogen		7	5

Szenario 3: Entkopplung nach dem Betriebsmodell, wobei die Milchprämie nicht an die Quote gebunden ist.

Szenario 4: Entkopplung nach dem Regionalmodell, wobei eine einheitliche regionale Prämie für Acker- und Grünland für das Jahr 2013 angenommen wird, wobei die Milchprämie dauerhaft als Betriebsprämie ausgezahlt wird.

Szenario 5: Entkopplung nach dem Regionalmodell mit regional einheitlicher Flächenprämie, wobei die Milchprämie ebenfalls in die regionale einheitliche Flächenprämie eingeht. Dieses entspricht dem deutschen Modell in der Endausgestaltung.

Szenario 6: Entkopplung nach dem Regionalmodell mit regional unterschiedlicher Acker- und Grünlandprämie, wobei die Milchprämie dauerhaft als Betriebsprämie ausgezahlt wird.

Szenario 7: Entkopplung nach dem Regionalmodell mit regional einheitlicher Flächenprämie, wobei die Milchprämie in die regionale Grünlandprämie eingeht.

Da Szenario 2 und 3 bis auf die Quotenrente identisch sind, wird im Folgenden nur noch Szenario 3 analysiert. Die jeweiligen Prämienzahlungen sind für die einzelnen Entkopplungsszenarien in Tabelle 4.8 aufgeführt.

Tabelle 4.8: Prämienhöhe bei unterschiedlichen Entkopplungsszenarien

		Szenario 4	Szenario 5	Szenario 6	Szenario 7
Schleswig-Holstein	Acker	276	359	332	332
	Grünland	276	359	186	405
Mecklenburg-Vorpommern	Acker	283	319	312	312
	Grünland	283	319	170	343

4.3.3 Veränderte ökonomische Rahmenbedingungen

Um die Auswirkung veränderter ökonomischer Rahmenbedingungen zu analysieren, wurden für das wahrscheinlichste Szenario 5 zusätzlich die in Tabelle 4.9 aufgeführten unterschiedlichen Output- und Inputpreisrelationen angenommen.

Tabelle 4.9: Szenarien für ökonomische Rahmenbedingungen des MTR

	niedrig	mittel	hoch
Lohnsatz zugekaufte AK [€/h]	10 (Base-run)	15 (Sze 55)	18 (Sze 65)
Opportunitätskosten Familien-AK [€/h]	3 (Base-run)	7,5 (Sze 75)	10 (Sze 85)
Preise der Grandes-Cultures (inkl. MWSt)	-10% (Sze 45)	±0% (Base-run)	+10% (Sze 35)

Dabei ist das Szenario eines um 10% erhöhten Getreidepreises äquivalent auch als ein Szenario eines Ertragszuwachses um 10% zu interpretieren. Es sei an dieser Stelle noch einmal explizit betont, dass die Ausgestaltung der Entkopplung keine direkten Auswirkungen auf die betriebliche Produktionsstruktur hat. Insofern ergeben sich für alle Entkopplungsszenarien grundsätzlich die gleichen komparativ-statischen Produktionseffekte, d.h. der in der Studie für Szenario 5 ermittelte Einfluss veränderter Input-Output-Preisrelationen auf die Produktionsstruktur gilt analog für die anderen Entkopplungsszenarien. Lediglich der Gewinn der Betriebe variiert über die jeweiligen Entkopplungsszenarien und damit ihre langfristige (dynamische) relative Wettbewerbsfähigkeit, so dass die Ausgestaltung der Entkopplung durchaus langfristig einen Einfluss auf die betriebliche Strukturentwicklung hat.

Literaturverzeichnis

Balmann, A., Lotze, H., and Noleppa, S. (1998). Agrarsektormodellierung auf der Basis typischer Betriebe, Teil 1. *Agrarwirtschaft*, 47(5):222–230.

Berkelaar, M. and Buttrey, S. (2004). *lpSolve*. R package version 1.0.1, http://cran.r-project.org.

Berkelaar, M., Eikland, K., and Notebaert, P. (2003). lp_solve, version 4.0. http://groups.yahoo.com/group/lp_solve/.

Bivand, R., Leisch, F., and Mächler, M. (2004). *pixmap: Bitmap Images*. R package version 0.4-2, http://cran.r-project.org.

Britz, W., Dominguez, I., Jansson, T., and Wieck, C. (2003). Auswirkungen der Mid-Term Review Beschlüsse auf die Landnutzung in europäischen Regionen und Betrieben — Eine quantitative Analyse mit dem CAPRI Modellsystem. Paper presented at the GeWiSoLa-Tagung in Hohenheim, 29.09.–01.10.2003.

Britz, W., Dominguez, I., Jansson, T., and Wieck, C. (2004). Model description. Zur Verfügung gestellt von der Universität Bonn.

Bruinsma, J., editor (2003). *World Agriculture: Towards 2015/2030. An FAO Perspective.* Earthscan Publications Ltd, London.

Henningsen, A. (2003). *linprog: Linear Programming / Optimization.* R package version 0.5-3, http://cran.r-project.org.

Institute for Agricultural Policy, University Bonn, Department of Economics, University College Galway, Institut Agronomique Mediterraneen de Montpellier, Departamento de Economia, Sociologia y Politica Agraria, Universidad Politecnica de Valencia, and Universita Degli Studi di Bologna, Dipartimento di Protezione e Valorizzazione Agro-Alimentare (Diproval) Sezione Economia (1999). Fair3 – ct96 – 1849; common agricultural policy regional impact analysis, final report. http://www.agp.uni-bonn.de/agpo/rsrch/capri/finrep.pdf.

Jansson, T., Junker, F., Pérez, I., and Wieck, C. (2003). Policy Description in the CAPRI Model. Working Paper 03-03, Institut für Agrarpolitik der Universität Bonn.

Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL) (2002). Taschenbuch Landwirtschaft 2002/03.

Landesanstalt für Landwirtschaft des Landes Brandenburg (2001). Datensammlung für die Betriebsplanung und die betriebswirtschaftliche Bewertung landwirtschaftlicher Produktionsverfahren im Land Brandenburg. 3. überarbeitete Auflage.

Landwirtschaftskammer Schleswig-Holstein (2002a). Rinder-Report 2001. Ergebnisse der Rinderspezialberatung in Schleswig-Holstein. Betriebswirtschaftliche Mitteilungen Nr. 554/555.

Landwirtschaftskammer Schleswig-Holstein (2002b). Rose Kälbermast.

Landwirtschaftskammer Schleswig-Holstein (2003a). Kalkulationsdaten. Mai 2003.

Landwirtschaftskammer Schleswig-Holstein (2003b). Rinder-Report 2002. Betriebswirtschaftliche Mitteilungen Nr. 568/2003.

Landwirtschaftskammer Schleswig-Holstein (2003c). Schweine-Report 2002. Ergebnisse aus der Schweinespezialberatung in Schleswig-Holstein. Betriebswirtschaftliche Mitteilungen Nr. 564/565.

Lecoutre, E. (2004). *R2HTML: HTML Exportation for R Objects.* R package version 1.4-3, http://cran.r-project.org.

Lewin-Koh, N. J. and Bivand, R. (2004). *maptools: Tools for Reading and Handling Shapefiles.* R package version 0.4-3, http://cran.r-project.org.

Neuwirth, E. (2004). *RColorBrewer: ColorBrewer Palettes.* R package version 0.2-1, http://cran.r-project.org.

R Development Core Team (2005). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Stabler, B. (2003). *shapefiles: Read and Write ESRI Shapefiles.* R package version 0.3, http://cran.r-project.org.

Kapitel 5

Economic Impact of the Mid-Term Review on Agricultural Production, Farm Income and Farm Survival: A Quantitative Analysis for Local Sub-Regions of Schleswig-Holstein in Germany

Arne Henningsen, Christian H.C.A. Henning, Carsten Struve und Joerg Mueller-Scheessel

erschienen in: Filippo Arfini (Hrsg.), Modelling Agricultural Policies: State of the Art and New Challanges, Proceedings of the 89th European Seminar of the European Association of Agricultural Economists (EAAE), Monte Università Parma, 2005, S. 757–772.

Abstract

This study analyzes the impact of the Mid-Term Review (MTR) on the agricultural sector in Schleswig-Holstein, a federal state in Germany. First, a very detailed farm group linear programming model is built to quantify the effects on agricultural production and farm incomes. The production adjustment to the MTR and its impact on farm profit vary significantly between individual farms. These results depend mainly on the farm type and the resource endowments of the farms. Second, the impact on structural change is examined with a farm survival model. Although the MTR clearly reduces the incomes of several farm types, it accelerates the structural change only gradually.

Keywords: policy reform, modeling production adjustment, farm income, structural change

5.1 Introduction

The Mid-Term Review (MTR) is certainly one of the most important reforms since the establishment of the Common Agricultural Policy (CAP). It is controversially discussed in particular, because it induces high political uncertainty for at least two reasons. First, the MTR contains new agricultural policy instruments, i.e. decoupling and cross compliance, for which their specific economic implications are not fully understood, yet. Second, in contrast to former CAP reforms the MTR includes a large range of policy options to be decided at national level, i.e. arrangement of decoupled payments. In this regard, farmers fear that depending on the decoupling option finally implemented at national level large income reductions and income redistribution among farm types will occur, while local politicians fear that agricultural production and employment will totally break down in specific local areas due to the reform. Finally, agricultural economists doubt that the MTR is really an effective political solution to the persisting structural adjustment problem in the agricultural sector.

Therefore, a detailed quantitative analysis of the economic impact of the MTR on agricultural production, farm income and farm survival is needed to reduce existing uncertainty and to provide a solid basis for rational evaluation of different reform scenarios.

In this regard the paper presents the results of a quantitative simulation analysis of the economic impact of the MTR on farm production, farm income and farm survival in 22 sub-regions of Schleswig-Holstein in Northern Germany.

In the following section a short outline of the agricultural sector in Schleswig-Holstein is given. The third section describes the model used to analyze the impact of the MTR. The model results are discussed in the fourth section. Finally, in the fifth section the paper is summarized and some conclusions are presented.

5.2 Agricultural Sector in Schleswig-Holstein

Schleswig-Holstein is Germany's most northern federal state (see figure 5.1). Its agricultural sector is relatively important with a share of 2.1% in total value added. This is twice as high as in overall Germany. Furthermore, agricultural productivity is one of the highest in Europe, especially for grain and milk production. Due to high yield of grain the previous area payments for grandes cultures amount to 429 €/ha, which is the highest in Germany.

Figure 5.1: Schleswig-Holstein in Germany

The average farm in Schleswig-Holstein has 55 ha agricultural land. Although this is comparably large in the "old" federal states of Germany, many farms in Schleswig-Holstein are too small to take full advantage of economies of scale.

Agricultural production is very heterogeneous in Schleswig-Holstein (see table 5.1). The main reason for this heterogeneity is the existence of different soils. Thus, it is convenient to divide Schleswig-Holstein into three main regions depending on the soil: "Marsch", "Geest" and "Hügelland" (see figure 5.2).

The "Marsch" is the most western part of Schleswig-Holstein on the coast of the North Sea. About half of the agricultural land in the "Marsch" is arable (53%). The clayey soils are highly productive, but they are also difficult and costly to cultivate. The arable land is

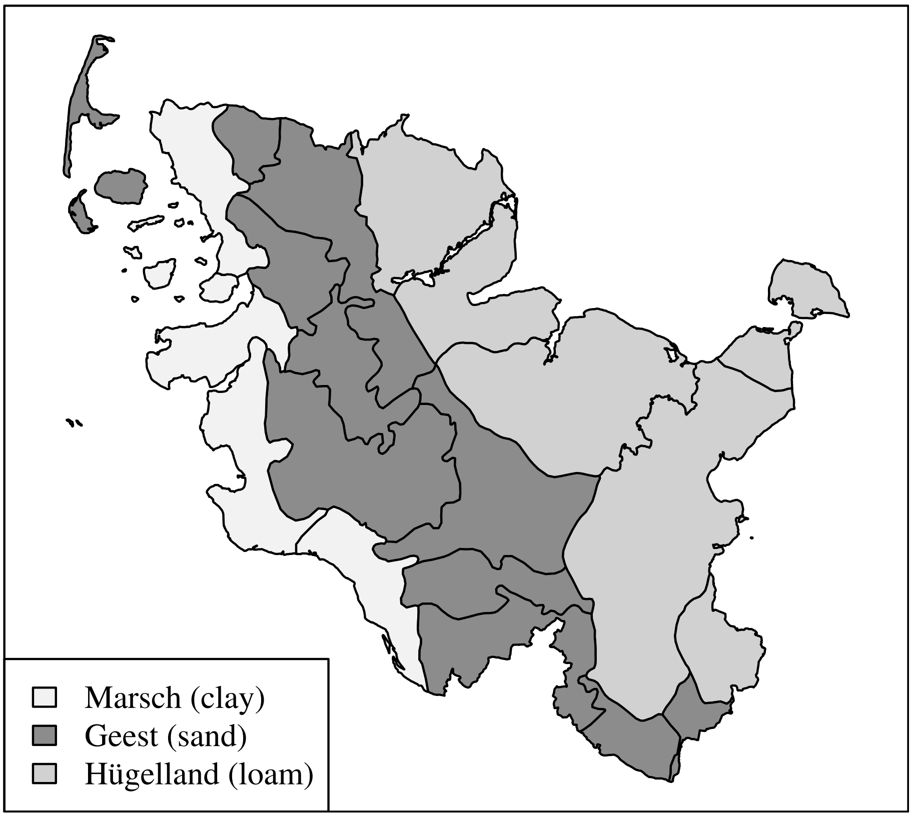


Figure 5.2: Regions and Soils in Schleswig-Holstein

Table 5.1: Regional Crop Areas (average 2000-2002)

	Marsch		Geest		Hügelland		Total
	1000 ha	%	1000 ha	%	1000 ha	%	1000 ha
Arable land	83	53	196	45	338	81	617
Cereals	54	34	76	18	197	47	327
Wheat	49	31	22	5	131	31	202
Rye + Triticale	1	1	30	7	16	4	47
Rapeseed	9	6	16	4	70	17	95
Feed prod. on arable land	4	3	81	19	37	9	122
Silage maize	2	1	59	14	20	5	81
Grass on arable land	2	1	21	5	16	4	39
Permanent grassland	74	47	236	55	80	19	390
Total agricultural land	157	100	432	100	418	100	1007

mainly used for wheat production, which generates very high yields on this soil. The grassland is mainly used as pasture for sheep, suckler cows and cattle fattening. Animal production is very unevenly distributed. Suckler cows, cattle fattening and pigs are concentrated in different subregions.

The "Geest" mainly lies in the center of Schleswig-Holstein. Most soils are more or less sandy, but there are also several bogs. Less than half of the agricultural land in the "Geest" is arable land (45%), because the bogs and several other areas are suitable only for grassland. The arable land is mainly cultivated with silage maize and cereals. While maize and rye grow quite well even on poor sandy soils, some areas with better soils are even capable to grow more demanding grains (e.g. barley, wheat) and rapeseed. Almost 10% of the arable land is used for grass production. Milk production is predominant in almost all subregions (on average 3420 kg milk quota per ha of total agricultural land). In several subregions there are also suckler cows. Intensive bull fattening based on maize is important especially in the northern subregions of the "Geest".

The "Hügelland" lies on the east coast adjacent to the Baltic Sea. The loamy soils are productive for many agricultural activities. Most of the agricultural land is arable (81%). The arable land is mainly used for wheat and rapeseed production. While cattle keeping is not very important in this region, there are several subregions with extensive pig production.

5.3 Model Description

The quantitative effects of the MTR on the agricultural sector in Schleswig-Holstein are analyzed with a sector model including 416 individual linear programming (LP) models. Each LP model corresponds to a specific farm size, farm type and subregion. The gross margin is maximized under general conditions of the years 2001 to 2003 (before MTR, "baserun") as well as under conditions that are expected in 2013 (after completion of MTR). The political conditions expected in 2013 can be simulated to analyze the effects of different national options like regionalization and partial decoupling. Furthermore, also economic conditions like prices can be simulated to analyze their impact on the agricultural sector. To take different land qualities, production intensities, crop rotation restrictions and cross compliance requirements into account, the LP model is strongly disaggregated. Each individual model includes roughly 1100 production activities and roughly 550 restrictions. Since a complete description of all details of the model would go beyond the scope of this article, only the most important features are described in this section. A more detailed description is available in Henning et al. (2004).

5.3.1 Farm-level Data

The arrangement of the 416 individual farm types is based on data that the farms submit when they apply for direct payments. Since almost all relevant farms receive some kind of direct payment, the data almost completely represent the agricultural sector of Schleswig-Holstein.

For each of these about 15,000 farms the data include the

- subregion, where the farm is located,
- agricultural area, divided into arable land and permanent grassland,
- area that is eligible for compensation payments for "grandes cultures",
- area used for sugar beet, potato and vegetable production,
- milk quota,
- received suckler cow premiums,
- received slaughter premiums,
- received special premium for male animals, divided into bulls and steers, and
- received ewe premiums.

The number of kept suckler cows and ewes as well as the number of slaughtered bulls, steers and other cattle can be directly calculated from the amounts of received premiums. The number of *kept* bulls, steers, dairy cows and other cattle is evaluated using additionally production and bookkeeping data (Landwirtschaftlicher Buchführungsverband, 2002a).

Since it is not possible with standard computers to calculate the LP model for all 15,000 farms in an appropriate time, we decided to construct a farm group model. Thus, the farms are divided into different groups according to the following attributes:

- 22 subregions
- 5 farm types, and
- 4 farm sizes.

The 22 subregions are selected to have relatively homogeneous soils and climatic conditions. The "Marsch" is divided into 5 subregions, the "Geest" is split into 11 subregions and the "Hügelland" is broken down into 6 subregions.

The farms are classified into farm types according to the proportions of the standard gross margins of the individual production areas. The five farm types used in the LP model are:

- specialized cash crop farms
- specialized forage-growing farms with predominantly milk production
- specialized forage-growing farms with predominantly beef production (cattle fattening and suckler cows)
- cash crop farms with some forage growing
- forage-growing farms with some cash crop production

The farm sizes are 5-60 ha, 60-100 ha, 100-200 ha and ≥200 ha agricultural land. Farms with less than 5 ha agricultural land are excluded from the model, because their total agricultural land accounts only for a small share in Schleswig-Holstein. Furthermore, these farms are either hobby farms or they produce special or niche products, so that their production decisions are hardly affected by the MTR.

From the 440 possible groups (22 subregions × 5 farm types × 4 farm sizes), 290 groups were selected, whose agricultural land accounts for at least 2% of the subregion or 0.05% of Schleswig-Holstein. For each group we calculated average values, to represent the group by an average farm in the LP model. Thus, the result for a total group can be calculated by multiplying the result of the individual (average) farm by the number of farms in the corresponding group. Furthermore, the aggregated results for each subregion, region or total Schleswig-Holstein can be identified by adding up the results of all groups in the respective area.

Unfortunately, the individual farm data provide no information on pig farming. Although pig production is not affected by the MTR, it has impact on other branches of production (e.g. via manure, labor requirements). Since this impact is only indirect, an approximate treatment of the pig production seems to be warrantable. Since most pigs are kept by cash crop farms, we assume for simplicity that all pigs are kept by farms that are so far considered as “specialized cash crop farms”. To implement this, we split this farm type into four farm types. All sizes of the “specialized cash crop farms” are divided into

- specialized cash crop farms without pig production
- cash crop farms with pig production
- pig farms with cash crop production
- specialized pig farms with some cash crop production

These subdivisions are done on a subregional level according to the proportion of these four farm types that are taken from a report based on bookkeeping data (Landwirtschaftlicher

Buchführungsverband, 2002b). The number of sows and fattening pigs per farm are taken from the same source. Due to the addition of farm types the number of farm groups increased from 290 to 416.

The area of permanent grassland was divided into permanent grassland on organic soils (bogs) and on mineral soils (sand, loam, clay). The particular proportions in each subregion are assessed by experts. Total agricultural area on mineral soils (all arable land plus permanent grassland on mineral soils) was split into ten quality categories. For each farm (group) the proportions of these categories are set equal to the proportions in the respective subregion. The permant grassland on organic soils (bogs) was divided into 5 quality categories. The proportions of each category are assessed by experts.

The data do not provide any information on family workers on the farms. Since the low opportunity costs of family workers on many farms heavily influence production decisions, this has to be evaluated (see table 5.2). We assume that there are 1.5 family workers on each dairy farm ("specialized forage-growing farms with predominantly milk production") and one family worker on each farm of other farm types. Further, we presume that each family worker works 2000 hours per year. One exception are "specialized cash crop farms without pig production". Since a farmer on this farm type cannot utilize his labour capacity during the winter, it is assumed that he can only work 1500 hours per year on the farm.

Table 5.2: Family Workers

Farm type	Family workers	Working hours
Specialized forage-growing farms with predominantly milk production	1.5	3000
Specialized forage-growing farms with predominantly beef production	1	2000
Cash crop farms with some forage-growing	1	2000
Forage-growing farms with some cash crop production	1	2000
Specialized cash crop farms without pig production	1	1500
Cash crop farms with pig production	1	2000
Pig farms with cash crop production	1	2000
Specialized pig farms with some cash crop production	1	2000

5.3.2 Activities and Restrictions

Each farm can choose its activities from roughly 1100 available possibilities. However, this choice is subject to roughly 550 restrictions. Data about the production activities are based on evaluations of bookkeeping data of farms in Schleswig-Holstein, data collections for planning purposes and assessments of experts. Furthermore, these data have been adjusted to converge the model results of the baserun with the real data. Depending on the availability of the data this calibration is done for subregions, regions or the whole federal state.

The farms can choose from all land cultivation activities that are relevant in Schleswig-Holstein. There are three main groups: cash crops, set-aside and forage production. The cash crops consist of several types of cereals, rapeseed, legumes, sugar beets, potatoes and cabbage. Set-aside activities include continuous fallow, rotational fallow, phacelia and non-food rapeseed. Forage production comprises silage maize, grass silage and pastures. Grass can be grown on arable land, permant grassland on mineral soils and on organic soils. The grass can be mowed once, twice, thrice or four times a year. If it is mowed less than four times a year, it can be used as pasture afterwards. Pastures can be cultivated with four different production intensities. All these production activities differ depending on the soil quality. Production activities on arable land additionally differ depending on the previous cropping on the same field. Especially the yield of wheat is heavily affected by the previous cropping and the yield of rapeseed decreases with an increasing share of cruciferous plants in the crop rotation.

The fertilization of the crops can be done by purchased mineral fertilizers as well as by manure. Leaching of nitrogen and potash is considered and depends on the soil quality and the kind of the fertilizer (mineral fertilizers or manure).

In our "realistic" scenario for 2013 we assume that the prices of cash crops do not change from 2003 to 2013. One exception is rye. The price of rye is assumed to decrease, because the MTR implies a discontinuation of the intervention of rye. Farms that feed their pigs with self-produced barley, rye or triticale benefit from the difference between the market price and the feeding value of these crops. Since the consumed quantities of potatoes and cabbage are more or less fixed, we assume in the model that the farmers cannot extend their sales quantity. At the time when the LP model was built, it was totally unclear how the European sugar market regime will change. Therefore, we also fixed the sugar beet production to the current amount.

The farms can choose from several kinds of animal production. These are dairy production, suckler cows, fattening of calves, intensive fattening of bulls with maize silage, fattening of bulls on pastures, sheep farming, sow keeping (farrow production) and pig fattening.

To account for economies of scale in dairy farming, the labor requirements per dairy cow decrease with an increasing herd size (see figure 5.3). This is implemented in the following way: The first dairy-cow requires 723 working hours, while all following cows require only 23 working hours.

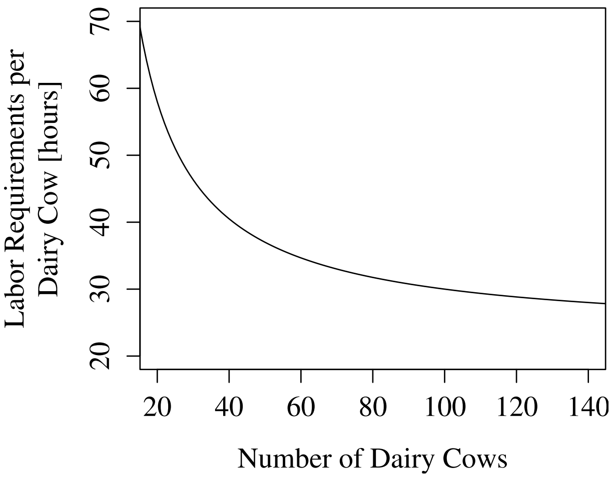


Figure 5.3: Labor Requirements per Dairy Cow

The milk production per cow of each farm was set equal to the average milk yield in the respective subregion. The feed requirements as well as the variable costs (e.g. concentrated feed, veterinary costs) increase with the milk yield per cow. It is assumed that the milk yield per cow increases by 10% from 2003 to 2013.

The requirements for forage of the cattle and sheep was divided into five parts: silage and pasture feed split into four grazing periods. While silage can be substituted for pasture feed, pasture feed of a certain grazing period cannot replace any other part of forage demand.

From the number of all kept animals, the indoor space for livestock husbandry was derived. It is assumed that 97.5% of the capacity for cattle and 100% of the capacity for sows and fattening pigs is utilized.

In the realistic scenario for 2013 we assume that the prices of beef, pork and sheep meat do not change from 2003 to 2013. However, due to the decoupling of the slaughter premium and the special premium for bulls, the prices of male calves will decrease. In the model the prices of bull calves are endogenous. They are chosen to get a market equilibrium in Schleswig-Holstein. This allows us to model also the impact of the decoupling on this price. The producer price of milk is assumed to decline from 0.291 €/kg in 2003 to 0.219 €/kg in 2013 due to the reduction of the intervention prices of butter and skimmed milk powder. In 2003 the surcharge for fat and protein was about 5% of the price in Schleswig-Holstein. We suppose that this surcharge will decrease proportional to the milk price and, thus, will remain 5% of the milk price.

The model contains all relevant premium payments. These premium schemes are exactly implemented in the model, both before the MTR as well as after the MTR. Since many forage-growing farms have only a small area with grandes cultures, also the small farmers scheme was implemented, again, both in the baserun and after the MTR. The modulation was 1% in the baserun and 5% in 2013.

The farms can hire labor as well as sell family labor. The wage for hiring labor is set to 10 €/hour. The opportunity cost of the family workers strongly varies. While it may be quite high for young well educated family members, it may be close to zero especially for older family members, who have almost no chance on the labor market. Furthermore the feasible wage outside the farm must be reduced by a certain amount to account for the preferences of most family members to work on their own farm. Due to these considerations we set the opportunity costs of the family workers in the model to 3 €/hour. In the long run (>15 years) this value would of course be higher.

5.3.3 Calculation of Profits

The calculation starts with the gross margin already maximized in the LP model. First, the profits of the farms are reduced by general expenses that are not accounted for in the LP model. These general expenses are taken from farm type specific bookkeeping data (Landwirtschaftlicher Buchführungsverband, 2002b).

Second, expenses for the tenancy of land and milk quota have to be deducted. These depend on the share of rented land and milk quota as well as on their price. Roughly half of the land and about 25% of the milk quota are rented in Schleswig-Holstein.

The shares of rented land (differentiated between arable land and permanent grassland) and milk quota represent statistical data and expert information. It is regarded that larger farms have higher shares of rented land than smaller farms.

The price for rented land depends on the shadow price of the specific soil category in the specific subregion and on the decoupled payments. The aggregated shadow price is calculated by taking the weighted average of the shadow prices of the particular soil category of all individual farms in the respective subregion. The area payments influence the price by the level of the payment and by the way these payments are established. To calculate the effect of decoupling we follow Isermeyer (2003). His model is extended to account for heterogeneous payments which occur either by single farm payments or by regional payments that differ between permanent grassland and arable land (for more details see Henning et al., 2004). Following his approach we presume that in case of single farm payments the payment entitlements are scarcer than land and, thus, land owners compete for entitlements. As a result the whole rent is transferred to the entitlement owners and the area payments are not included in the rental prices, which correspond only to the average shadow prices in the respective subregion.

In case of regional payments we again follow Isermeyer (2003) and assume that land is relatively rare compared to the payment entitlements. Then the competition for land will raise the willingness to pay for land by the level of the regional payments. Thus, in this case the rental prices correspond to the average shadow prices plus the level of the regional payments.

5.3.4 Implementation of the Model

The programming of the model and the analysis of the model results are implemented in the free language and environment for statistical computing "R" (R Development Core Team (2005), see also http://www.r-project.org). The underlying linear programming models have been solved with the R packages "lpSolve" (Berkelaar et al., 2005) and "linprog" (Henningsen, 2003) that internally use the LP software "lp_solve" (Berkelaar et al., 2003). The maps are produced with the R packages "shapefiles" (Stabler, 2003) and "maptools" (Lewin-Koh and Bivand, 2004).

5.3.5 Analysis of the Structural Change

Though the LP model described above is certainly appropriate to analyze the short and medium-term effects of the MTR on production decisions and farm income, it cannot examine the long-term effects on structural change.

Farm structure developments are determined by survival of farms. We divide farms into different categories according to farm type and size. The effects of the MTR on the survival of different farm categories are analyzed using following model:

$$N_i^t = N_i^0 \left(p_i^t q_i + \left(1 - p_i^t\right) r_i^t + s_i^t \right) \tag{5.1}$$

where for each farm category i, N_i^0 is the number of farms in the base period, N_i^t is the number of farms at time t, p_i^t is the share of farmers that retire between the base period and time t, q_i is the share of retired farmers who have a successor, $(1 - r_i^t)$ is the share of farms that leave the farm category between the base period and time t although the farmer does not retire (e.g. switch to another farm category), and s_i^t is the number of farms that enter the farm category between the base period and time t represented as a share of N_i^0.

According to empirical studies farm survival in Germany is mainly determined via farm succession (for a literature overview see Tietje, 2004, chapter 4). Thus, q_i is the most important factor influencing structural change. To determine the share of farms that have a successor (q_i) we use an existing model on farm succession decisions estimated for farms in Schleswig-Holstein (Tietje, 2004). This model estimates the probability of succession for individual farms as a function of farm type, size, profit and other socio-economic determinants. Based on these results the probability of succession can be calculated for each farm category and for different levels of profit.

Our model on the structural change (equation (5.1)) is calibrated using data of the agricultural censuses of 1991 and 1999. The year 1991 is taken as base period and year 1999 as time t. For each farm category the number of farms in 1991 (N_i^0) and 1999 (N_i^t) as well as the share of farmers who retired between 1991 and 1999 (p_i^t) are taken from these data. The share of retired farmers who have a successor (q_i) is evaluated for each farm category using the model of Tietje (2004). Thus, only r_i^t and s_i^t are unknown, and assuming reasonable values for s_i^t, r_i^t can be directly calculated.

Having the model on farm survival (5.1) fully specified, we use it to forecast the development of the farm structure. Assuming an exponential growth model ($N_i^t = N_i^0 e^{w_i t}$) the annual growth rates (w_i) of each farm category can be calculated by

$$w_i = \frac{\ln\left(p_i^t q_i + \left(1 - p_i^t\right) r_i^t + s_i^t\right)}{t} \tag{5.2}$$

where t, the time period between the two agricultural censuses, is 8 in our case. As the impact of MTR on farm incomes is already known from the LP model, the share of retired farmers who have a successor (q_i) can be additionally evaluated with farm incomes after the MTR. Thus, assuming that p_i^t, r_i^t and s_i^t do not change, the annual growth rates of different farm categories after the MTR can be calculated using equation (5.2).

5.4 Results

5.4.1 Realistic Scenario

At first the results of the realistic scenario are presented. This scenario assumes full decoupling where decoupled payments are introduced as uniform regional premiums, which will have reached 359 €/ha in 2013. Furthermore, the prices of most crop products and meat do not change, but the price of rye is slightly reduced and the milk price is clearly reduced (see section 5.3).

The optimal adjustment of farm production to MTR until 2013 varies significantly over individual farms. In the short and medium term individual resource endowments (i.e. milk quota, stable capacities, land and soil quality) are the most important determinants of the adjustment at individual farm level.

However, on average specific adjustment patterns can be observed for different farm types. Significant adjustments can be observed for forage-growing farms and mixed farms, while cash crop and pig farms do not significantly adjust their output structure to the MTR. On average forage-growing farms reduce bull fattening by 22% and suckler cows by 98%. Specialized dairy farms continue to produce milk and fully use their milk quota despite the milk price reduction. All forage-growing farms extensify their forage production. While forage-growing farms on low quality soils (i.e. especially in the "Geest" region) do this by increasing the area for forage production and reducing cash crop production, forage-growing farms on better soils (i.e. "Marsch" region) reduce the number of animals. Moreover, specialized dairy farms and grain farms continue to use the large part of their land for agricultural production and cease production only on a small part of their land (max. 3%). In contrast, forage-growing farms specialized in bull fattening or suckler cow farming cease production on a significant share of their land ranging up to 45% in specific areas. Farms predominantly cease production on low quality permanent grassland.

Farm incomes are significantly reduced by 20% on average due to the MTR. However, the impact of the MTR on farm profits varies significantly over individual farms, farm types and regions. In particular, dairy farms observe on average the highest profit reductions ranging from -24% up to -37%, while the income of cash crop and pig farms is not much affected by the MTR (between -4% and +8%). Forage-growing farms specialized in beef production on average observe a rise in profit of 17%. For all farm types profit reductions are lower for small farms compared to large farms. This has several reasons. The most important causes are the following: First, the modulation of the payments favors small farms. Second, small cash crop and pig farms make a higher share of their profits with products that are not negatively affected by the MTR (e.g. pig farming, cabbage) than large farms of the same type. Third, the share of costs in total revenue is smaller for small farms than for large farms, because the latter have to hire workers. Thus, for farms that can only minimally adjust their production (i.e. cash crop, pig and specialized dairy farms) a proportional reduction of the revenues (e.g. milk sales or area payments) implies that the profits of large farms are relatively more reduced than the profits of small farms.

At aggregate level agricultural production adjustments are much more moderate when compared to individual farm level. On average the area of cash crops is reduced by 2%

and the area for forage growing decreases by 1%. However, one can observe clear regional adjustment patterns (see figure 5.4).

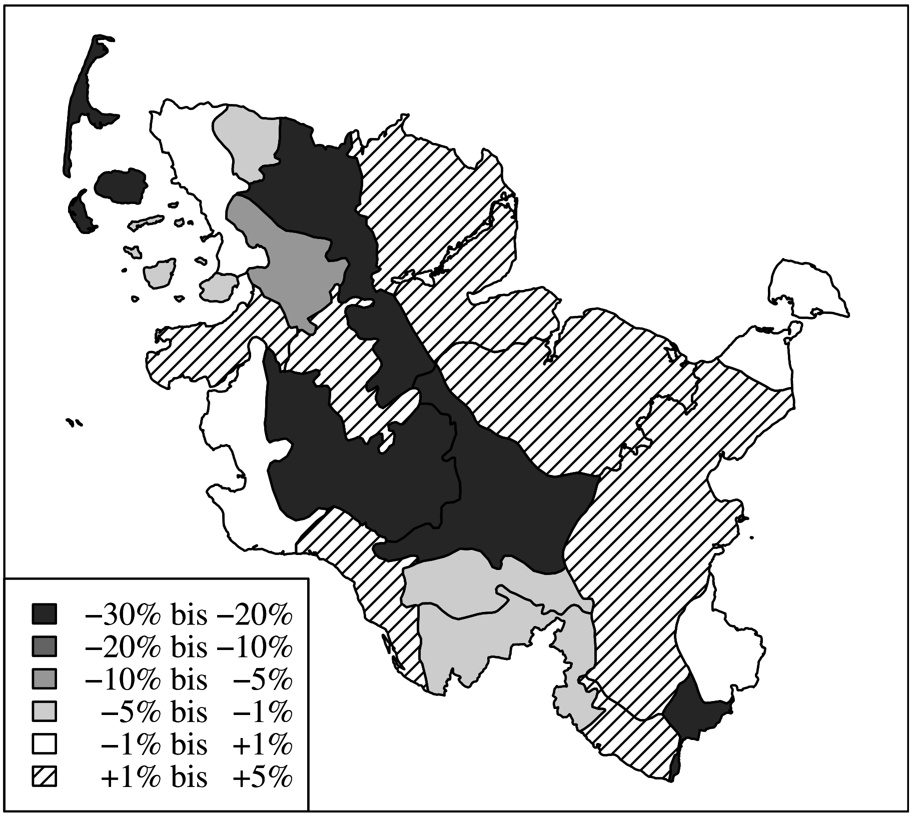


Figure 5.4: Changes of Cash Crop Production

In the "Marsch" and "Hügelland", where the soils are highly productive for grain farming, the cash crop production is slightly increased (+0.4% and +1.7%), while forage production is reduced (-1.2% and -5.6%, respectively). On the other hand, in the "Geest", where the soils are less productive, cash crop production is reduced by 16% and forage growing is increased by 1%.

About 1.5% of total agricultural land is no longer used for production and will only be maintained to receive the decoupled premium. This land is predominantly low-quality permanent grassland and its share of all agricultural land varies from 0.3% in the "Marsch" to 2.8% in the "Geest".

In spite of the strong decrease of the milk price the milk quota is still fully used. The number of dairy cows is reduced by 8%, because the milk yield per cow (+10%) rises more than the milk quota (+1.5%). Suckler cow farming is reduced in all regions by more than 90%. Due to the reduction of dairy and suckler cows less calves are born and, thus, less bulls can be fattened. Furthermore, the augmentation of calf fattening instead of bull fattening reduces the number of kept male cattle. The reduction of male cattle for fattening is on average 16%, but varies considerably between regions. It is 48% in the "Marsch", 11% in the "Geest" and only 1% in the "Hügelland". Due to the decoupling of the special premium for bulls and the slaughter premium the price of bull calves decreases by 77%.

The area payments for grandes cultures before the MTR and the decoupled premiums after the MTR are both shifted to the land owners via the rent for land. Thus, the decrease of the area payments from 429 €/ha to 359 €/ha is the main reason for the decline of the shadow prices of arable land (-17%).

Since the new decoupled area premiums for permanent grassland are shifted to the land owners to a larger extent than the previous animal premiums, the shadow prices of permanent grassland rise on average by 34%. Especially the shadow prices of low quality permanent grassland that are very low before the MTR rise to the level of the decoupled premium minus the costs to maintain the area (e.g. mowing).

The shadow price of milk quota is reduced by the decrease of the milk price and the quota enlargement. However, the decoupling of other cattle and crop premiums worsens the alternative utilization of land and labor which gives a positive impact on the quota value. On average the shadow price of the milk quota decreases by 39%.

5.4.2 Alternative Scenarios

In the following section a few results of alternative scenarios are presented.

Since decoupled payments do not affect production decisions, the distribution of these payments (e.g. single farm payments, regional uniform payments, regional payments differentiated between arable land and permanent grassland, milk premium farm specific or regionalized) does not affect production decisions. However, this of course strongly influences the income of individual farms and the shadow prices of land and milk quota.

While a unified regional decoupled premium (realistic scenario) reduces farm incomes on average by 20%, single farm payments do not reduce the average farm profit. The main reason for this is that single farm payments lead to much lower rents for agricultural land (see section 5.3.3).

Interestingly, individual farm profit developments and premium payments are not perfectly correlated across decoupling scenarios, because the decoupling scenarios not only affect the premium payments, but also the prices of land and milk quota. For example, dairy farms receive the highest premium payments assuming unified regional premium payments and a farm specific milk premium. However, these farms realize their lowest profit loss assuming single farm payments.

5.4.3 Structural Change

The model described in section 5.3.5 is used to project the structural change of the agricultural sector in Schleswig-Holstein. Furthermore, it is analyzed how the structural change is influenced by the MTR. Projected annual growth rates of different farm types and farm sizes with and without the MTR are shown in table 5.3.

Table 5.3: Annual Growth Rates of Different Farm Types

	without MTR	with MTR
Cash crop farms		
$\leq$ 50 ha	-4.5%	-4.6%
> 50 ha	-1.1%	-1.2%
Forage-growing farms		
$\leq$ 50 ha	-4.1%	-4.3%
> 50 ha	0.0%	-0.2%
All farms		
$\leq$ 50 ha	-4.2%	-4.4%
> 50 ha	-0.3%	-0.5%

The results show that although the MTR has a clear impact on farm profits in the short and medium run, induced profit reductions have only little impact on the survival of individual farms and, thus, on the long-run development of farm structure. Hence, structural change is independent of the MTR characterized by a clear decrease of small farms and an increasing average farm size. Our projections for 2030 show that the average farm size would be 94 ha without MTR and will be 100 ha with MTR.

5.5 Summary and Conclusions

The Mid-Term Review is one of the most important reforms since the establishment of the CAP. It contains new agricultural policy instruments, for which specific economic implications are not fully understood, yet. Therefore, the quantitative economic impact of the MTR on agricultural production, farm income and structural change is analyzed in this paper. This analysis is carried out exemplarily for Schleswig-Holstein, the most northern federal state of Germany. The agricultural sector of Schleswig-Holstein is highly productive and is characterized by very heterogeneous conditions for agricultural production. Especially the soil quality strongly differs between regions.

The effects of the MTR on agricultural production and farm income until 2013 are modeled using a detailed farm group linear programming (LP) model. The main advantage of this model is its excellent data base, because it is based on data of virtually all 15,000 farms in Schleswig-Holstein. These data provided by the department of agriculture are based on the data that farms submit when applying for direct payments.

Since farm survival is mainly determined via farm succession, we analyzed the effects of the MTR on structural change using an existing model on farm succession decisions estimated for farms in Schleswig-Holstein (Tietje, 2004). As the decision on farm succession depends also on the profits of the farm, we are able to analyze the effects of the MTR on structural change via the effects of the MTR on farm profits using the results of the LP model.

The effects of the MTR on optimal adjustment of farm production and on farm profits vary significantly over individual farms. Resource endowments are the most important determinants of the adjustment. While forage-growing farms and mixed farms significantly adjust to the MTR, cash crop and pig farms do not. Forage-growing farms reduce bull fattening by 22% and suckler cows by 98%. Dairy farms still fully use their milk quota despite the milk price reduction. All forage-growing farms extensify their forage production. The MTR reduces average farm incomes by 20%, but this varies significantly over farm types, farm sizes and regions. On average dairy farms observe the highest profit reductions (-31%).

At aggregate level agricultural production adjustments are much more moderate when compared to individual farm level. However, one can observe clear regional adjustment patterns. In the regions with good soils grain production slightly increases and forage growing slightly decreases. However, in regions with poor soils grain production significantly decreases and forage production slightly increases. Only a small proportion of land is no longer used for production. Though the MTR has a significant influence on farm income, its impact on the farm structure is rather low.

Although the MTR introduces new policy instruments, it does not significantly change aggregate agricultural production and farm structure - at least in Schleswig-Holstein. However, significant changes can be observed at regional and farm level. Since the effects of the MTR strongly depend on the resource endowments of the farms, disaggregated farm level models are necessary to analyze the specific impact of the MTR. Furthermore detailed farm level data are needed for modeling. Thus, the modeling approach and the data base used in this analysis are very suitable for modeling agricultural policies.

Bibliography

Berkelaar, M., Buttrey, S., and others (2005). *lpSolve: Interface to Lp_solve v. 5 to Solve Linear/Integer Programs*. R package version 1.1.9, http://cran.r-project.org.

Berkelaar, M., Eikland, K., and Notebaert, P. (2003). lp_solve, version 4.0. http://groups.yahoo.com/group/lp_solve/.

Henning, C. H. C. A., Henningsen, A., Struve, C., and Müller-Scheeßel, J. (2004). *Auswirkungen der Mid"=Term"=Review"=Beschlüsse auf den Agrarsektor und das Agribusiness in Schleswig"=Holstein und Mecklenburg"=Vorpommern.* Agrarwirtschaft, Sonderheft 178. Agrimedia, Bergen/Dumme.

Henningsen, A. (2003). *linprog: Linear Programming / Optimization.* R package version 0.5-3, http://cran.r-project.org.

Isermeyer, F. (2003). Umsetzung des Luxemburger Beschlusses zur EU-Agrarreform in Deutschland — eine erste Einschätzung. Arbeitsbericht 03/2003, Institut für Betriebswirtschaft, Agrarstruktur und ländliche Räume, Braunschweig.

Landwirtschaftlicher Buchführungsverband (2002a). Betriebszweigabrechnung 2001/2002.

Landwirtschaftlicher Buchführungsverband (2002b). Wirtschaftsergebnisse 2001/2002.

Lewin-Koh, N. J. and Bivand, R. (2004). *maptools: Tools for Reading and Handling Shapefiles.* R package version 0.4-3, http://cran.r-project.org.

R Development Core Team (2005). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Stabler, B. (2003). *shapefiles: Read and Write ESRI Shapefiles.* R package version 0.3, http://cran.r-project.org.

Tietje, H. (2004). *Hofnachfolge in Schleswig-Holstein.* PhD thesis, Christian-Albrechts-Universität, Kiel, Germany.

Kapitel 6

Analyse von Transaktionskosten auf dem ländlichen Kreditmarkt in Polen

Anja Wiebusch, Arne Henningsen und Christian H.C.A. Henning

erschienen in: Dabbert, S., W. Grosskopf, F. Heidhues und J. Zeddies (Hrsg.) (2004): Perspektiven in der Landnutzung — Regionen, Landschaften, Betriebe — Entscheidungsträger und Instrumente. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaus e.V., Band 39, Landwirtschaftsverlag, Münster-Hiltrup, S. 363–371.

6.1 Einleitung

Die Funktionsfähigkeit der ländlichen Kreditmärkte ist eine basische Voraussetzung für die Wettbewerbsfähigkeit der landwirtschaftlichen Produktion in den mittel- und osteuropäischen Transformationsländern. Erhält der landwirtschaftliche Produktionssektor nicht im angemessenen Maße Kredite, so wird er in der Entwicklung hinter den übrigen Sektoren der Volkswirtschaft hinterherhinken. Da in vielen mittel- und osteuropäischen Ländern der landwirtschaftliche Sektor noch eine verhältnismäßig bedeutsame gesamtwirtschaftliche Stellung hat und auch ein relativ großer Anteil der Erwerbsbevölkerung in der Landwirtschaft beschäftigt ist, hat die Funktionsfähigkeit der ländlichen Kreditmärkte auch eine große gesamtwirtschaftliche Bedeutung.[1] Die statistischen Daten deuten jedoch auf eine geringe Investitionstätigkeit in der polnischen Landwirtschaft hin; so betragen die landwirtschaftlichen Investitionen nur 113 Zl/ha[2], und ein polnischer Traktor ist durchschnittlich 19 Jahre alt (Instytut Ekonomiki Rolnictwa i Gospodarki Zywnosciowej (IERiGZ), 1998, S. 3). Der polnische Staat versucht deshalb, durch umfangreiche Subventionsprogramme die Zinsbelastung für die Landwirte zu senken, um die Investitionstätigkeit in der Landwirtschaft zu erhöhen.

In der Literatur wird der Zustand des ländlichen Kreditmarktes in Polen kontrovers diskutiert. Während für die einen die geringe Investitionstätigkeit in der Landwirtschaft auf einen Mangel an profitablen Investitionsprojekten zurückzuführen ist (Petrick, 2000), gehen andere davon aus, dass die geringe Investitionstätigkeit vielmehr an einem erschwerten Zugang zum ländlichen Kreditmarkt liegt (Moosburger et al., 1999, S. 360; Puslecki, 2000, S. 99). Als Hauptursachen für einen erschwerten Zugang der Landwirte zum Kreditmarkt werden zum einen die im Verhältnis zur Kreditsumme hohen Transaktionskosten und zum anderen das im Verhältnis zu anderen Wirtschaftssektoren höhere Risiko gesehen (Koester, 2001, S. 306f).

Ziel dieses Beitrages ist es, eine Methode vorzustellen, mit der man die Wirkungen der auf dem ländlichen Kreditmarkt vorherrschenden Transaktionskosten auf die Kapitalausstattung der landwirtschaftlichen Betriebe analysieren kann. Diese Methode wird dann auf Daten polnischer landwirtschaftlicher Betriebe und schleswig-holsteinischer Testbetriebe angewandt, um den Transaktionskosteneinfluss auf dem ländlichen Kreditmarkt vergleichen zu können.

[1] In Polen arbeiten 18,8% der Erwerbsbevölkerung in der Landwirtschaft.

[2] Zum Vergleich: Die landwirtschaftlichen Investitionen pro Hektar belaufen sich in Deutschland auf das 15fache und die Investitionen in die slowakische Landwirtschaft immerhin auf das 3fache der polnischen Investitionen.

6.2 Vorgehensweise

Um die Auswirkung der Transaktionskosten auf die Agrarkreditvergabe bzw. Kapitalausstattung der landwirtschaftlichen Betriebe analysieren zu können, wird wie folgt vorgegangen:

Zunächst wird auf der Grundlage von Befragungsdaten eine aggregierte Produktionsfunktion mit Kapital, Arbeit, Vorleistungen und Boden als Produktionsfaktoren geschätzt. Da es nicht sinnvoll ist, die Mengeneinheiten der unterschiedlichen landwirtschaftlichen Outputs für die Schätzung aufzusummieren, wird als endogene Größe der landwirtschaftliche Umsatz genommen:

$$U = f\,(K, L, V, B) \tag{6.1}$$

Wir haben eine quadratische Funktionsform[3] gewählt, somit ist die Produktionsfunktion wie folgt spezifiziert:

$$U = \alpha + \sum_{i \in \mathrm{PG}} \beta_i x_i + \frac{1}{2} \sum_{i \in \mathrm{PG}} \sum_{j \in \mathrm{PG}} \gamma_{ij} x_i x_j \qquad \mathrm{PG} = \{K, L, B, V\} \tag{6.2}$$

wobei $\gamma_{ij} = \gamma_{ji}$, mit

U Umsatz pro Betrieb in Landeswährung

x_K Anlagevermögen (ohne Boden) pro Betrieb in Landeswährung

x_L Arbeitskräfte pro Betrieb

x_V Vorleistungen pro Betrieb in Landeswährung

x_B Boden pro Betrieb in ha

Ferner berechnet sich das Wertgrenzprodukt des Kapitals aus:

$$\mathrm{WGP}_K = \frac{\partial U}{\partial x_K} = \beta_K + \sum_{i \in \mathrm{PG}} \gamma_{Ki} x_i \quad \text{mit } \gamma_{ij} = \gamma_{ji} \tag{6.3}$$

Da im Optimum das Wertgrenzprodukt des Kapitals den Kapitalkosten, d.h. der Annuität A[4] entspricht, gilt:

$$\mathrm{WGP}_K = A = \frac{(i+1)^N \cdot i}{(i+1)^N - 1} \tag{6.4}$$

mit

[3]Vgl. Fuss und McFadden (1978, S. 238)

[4]Die Annuität spiegelt die Kapitalkosten (Zinsen und Tilgung) wieder.

A Annuität in Landeswährung

N Laufzeit in Jahren

i Zinssatz in Prozent

Vernachlässigt man zunächst die individuelle Risikoaversion, so stellt die Differenz i_{Diff} aus dem aus Gleichung (6.4) berechneten optimalen Zinssatz $\hat{i}$ und dem tatsächlichen realen Zinssatz i einen Indikator für einen unvollkommenen Kapitaleinsatz auf Mikroebene dar:

$$i_{\text{Diff}} = \hat{i} - i \tag{6.5}$$

Eine positive Zinsdifferenz zwischen $\hat{i}$ und tatsächlichem realem Zinssatz i deutet darauf hin, dass die landwirtschaftlichen Betriebe tendenziell unterkapitalisiert sind. Eine negative Zinsdifferenz lässt darauf schließen, dass die landwirtschaftlichen Betriebe tendenziell überkapitalisiert sind. Eine Zinsdifferenz von Null würde den Optimalfall widerspiegeln, dass der tatsächliche reale Zinssatz dem Schattenpreis des Kapitals entspricht. Dieser Optimalfall ist in der Realität nicht zu erwarten, da man nicht davon ausgehen kann, dass selbst bei funktionsfähigen Kreditmärkten immer optimal investiert wird. Deshalb ist vielmehr die Verteilung der Zinsdifferenz über die einzelnen landwirtschaftlichen Unternehmen von Bedeutung.

Berücksichtigt man hingegen die Risikoaversion landwirtschaftlicher Betriebe, so folgt, dass sich eine systematische positive Abweichung des Optimalzinssatzes $\hat{i}$ von dem zu zahlenden Realzinssatz i ergibt. Vereinfacht lässt sich die Risikoaversion mit Hilfe eines Risikoaufschlags λ messen, d.h. es gilt im mikroökonomischen Gleichgewicht:

$$\hat{i} = (1 + \lambda)\, i \tag{6.6}$$

Da die jeweilige Risikoaversion nicht bekannt ist, besteht die Möglichkeit, die Verteilung der Risikoaversion λ über die einzelnen Betriebe mit Hilfe einer Monte Carlo Simulation zu simulieren. Hierbei ist noch ein Störterm ε zu berücksichtigen, der beispielsweise witterungsbedingte Störungen einbezieht. Für den Realzinssatz i gilt:

$$i = \frac{\hat{i}}{1 + \lambda} \tag{6.7}$$

Durch das Einsetzen von Gleichung (6.7) in (6.5) und Berücksichtigung eines Störterms folgt:

$$i_{\text{Diff}} = \hat{i} - \frac{\hat{i}}{1 + \lambda} + \varepsilon = \hat{i}\left(1 - \frac{1}{1 + \lambda}\right) + \varepsilon \tag{6.8}$$

Die Risikoaversion λ kann dabei bei der Simulation nur Werte größer oder gleich Null annehmen, da eine Risikoaversion zu einem Risikoaufschlag und nicht -abschlag führt. Der Störterm ε hingegen kann sowohl positive als auch negative Werte annehmen.

Auf diese Weise lässt sich die durchschnittliche Verteilung der Differenz zwischen $\hat{i}$ und risikokorrigierten Zinsen unter der Annahme, dass keine Transaktionskosten oder asymmetrische Marktzugangsbarrieren bestehen, ableiten. Ein Vergleich der tatsächlichen empirischen Verteilung der Zinsdifferenz zwischen Optimalzinssatz $\hat{i}$ und tatsächlichem Zins i mit der Verteilung aus der Monte Carlo Simulation ermöglicht, systematische Rückschlüsse auf vorherrschende Transaktionskosten und Marktzugangsbarrieren zu ziehen. Ein Problem bei der Anwendung dieser Methode ist jedoch, dass wir den durchschnittlichen und nicht den marginalen Zinssatz für unsere Berechnungen ansetzen. Dies führt zu einer leichten Unterschätzung des tatsächlichen realen marginalen Zinssatzes.

6.3 Schätzung der Produktionsfunktion

Für die Schätzung der Produktionsfunktion wurden die Wirtschaftsdaten von 464 landwirtschaftlichen Betrieben unterschiedlicher Rechtsformen aus den früheren Woiwodschaften Szczecin, Tarnów und Rzeszów genutzt. Die Stichprobe wurde 1999 durch das IAMO erhoben und enthält sowohl die Wirtschaftsdaten von Buchführungsbetrieben als auch von Betrieben ohne Buchführung.

Die Schätzung von Gleichung (6.2) liefert nicht signifikante Schätzer, während der F-Test auf eine Signifikanz der Gesamtregression hindeutet. Der R^2-Wert liegt bei 0,75. Das Ergebnis deutet auf Multikollinearität der exogenen Variablen hin. Eine Korrelationsanalyse zeigt, dass die Produktionsfaktoren in sehr starkem Maße voneinander abhängig sind. Multikollinearität führt zwar zu unverzerrten Schätzern, die Varianz der Schätzer ist jedoch so groß, dass sie sehr weit vom tatsächlichen Wert entfernt liegen können. Berechnet man den Standardfehler für das Wertgrenzprodukt des Kapitals, so zeigt sich, dass dieser relativ groß ist. Die meisten Betriebe haben einen Standardfehler zwischen 0,05 und 0,1.[5]

[5] In der Literatur werden zwei statistische Verfahren genannt, die beim Vorliegen von Multikollinearität eingesetzt werden können: Die Hauptkomponentenanalyse (Principal Component Analysis) und die Ridge-Regression. Bei der Hauptkomponentenanalyse werden mit Hilfe der exogenen Variablen X hypothetische Variablen Z als Linearkombination der x_i konstruiert, die voneinander unabhängig sind. Der größte Nachteil der Hauptkomponentenanalyse ist, dass sich die einzelnen Komponenten häufig nicht sinnvoll durch die ursprünglichen exogenen Variablen erklären lassen (vgl. Cooper, 1990, S. 19). Da wir mit unserer Schätzung das Wertgrenzprodukt des Kapitals ermitteln wollen, ist diese Methode zur Eindämmung der Multikollinearität somit ungeeignet. Die Ridge-Regression ist eine weitere Möglichkeit, dem Multikollinearitätsproblem zu begegnen. Bei der Ridge-Regression werden (leichte) Verzerrungen an den Regressionsschätzern vorgenommen, um ein besseres Regressionsoutput für die Schätzer zu erhalten. Die-

Brooks (2002, S. 192f) weist darauf hin, dass ein hoher Grad an Multikollinearität auch häufig mehr ein Datenproblem als ein Problem des Modells ist. Er schlägt deshalb vor, eine der kollinearen Variablen aus der Schätzung zu entfernen, die Anzahl der Beobachtungen zu erhöhen, oder die Multikollinearität zu ignorieren, wenn das Modell ansonsten adäquat ist. Die Anwendung des Modells auf Daten schleswig-holsteinischer Betriebe zeigt jedoch, dass das Modell als adäquat anzusehen ist und die Multikollinearität eher auf ein Problem der Datengrundlage zurückzuführen ist. Im folgenden wird deshalb die Multikollinearität in Kauf genommen.

6.4 Testen auf Konkavität

Im folgenden ist zu testen, ob die Produktionsfunktion die Bedingung der Konkavität erfüllt.

Eine Produktionsfunktion ist konkav, wenn die Hessesche Matrix negativ semidefinit ist, d.h. wenn alle ungeraden Hauptminoren kleiner oder gleich Null und alle geraden Hauptminoren größer oder gleich Null sind. Eine andere Möglichkeit ist, die Eigenwerte zu berechnen. Für die Konkavität der Produktionsfunktion müssen alle Eigenwerte negativ sein, dies ist bei der vorliegenden Produktionsfunktion nicht gegeben.

6.5 Erzwingung von Konkavität

Die Konkavität der Produktionsfunktion läßt sich dadurch erzwingen, dass die Koeffizienten γ_{nn} der Hesseschen Matrix durch die Elemente der Cholesky-Zerlegung ersetzt werden. Die zu restringierende Matrix wird dabei folgendermaßen zerlegt:

$$D_{ij} = -KK' = -\begin{bmatrix} K_{11} & 0 & 0 & 0 \\ K_{21} & K_{22} & 0 & 0 \\ K_{31} & K_{32} & K_{33} & 0 \\ K_{41} & K_{42} & K_{43} & K_{44} \end{bmatrix} \begin{bmatrix} K_{11} & K_{21} & K_{31} & K_{41} \\ 0 & K_{22} & K_{32} & K_{42} \\ 0 & 0 & K_{33} & K_{43} \\ 0 & 0 & 0 & K_{44} \end{bmatrix} \tag{6.9}$$

Somit gilt für die einzelnen Elemente von D:

$$D_{ij} = -\sum_{k=1}^{j} K_{jk} K_{ik} \; \forall \; i \geq j \tag{6.10}$$

ses Vorgehen führt jedoch dazu, dass die Schätzer ebenfalls nicht mehr „blue“ sind, da man lediglich versucht, eine Eigenschaft der Schätzer zu verbessern, indem man eine andere Eigenschaft verschlechtert.

Die Koeffizienten γ_{ij} in der Gleichung (6.2) werden nun durch D_{ij} substituiert:

$$\gamma_{ij} = D_{ij} \tag{6.11}$$

Die nicht-linear Schätzung der modifizierten Gleichung (6.2) konvergiert jedoch nicht.

6.6 Zweistufige Minimum-Distance Schätzung

Eine weitere Möglichkeit, die Konkavität zu erzwingen, ist eine zweistufige Minimum-Distance Schätzung. Hierbei werden die Parameter in zwei Stufen geschätzt. Im ersten Schritt wird ein unrestringiertes Modell geschätzt. In einem zweiten Schritt werden dann die restringierten Koeffizienten, die die Konkavitätsbedingung erfüllen (siehe Gleichung (6.10)), mit Hilfe einer Minimum-Distance Schätzung generiert (vgl. Koebel, 1998; Koebel et al., 2003). Anstelle der nicht restringierten Koeffizienten $\hat{\gamma}_{ij}$ sollen restringierte Koeffizienten $\hat{\gamma}^0_{ij}$ nach Gleichung (6.9) gefunden werden, die die Konkavitätsbedingung erfüllen und gleichzeitig sehr nah an den unrestringierten Koeffizienten $\hat{\gamma}_{ij}$ liegen. Die Minimum-Distance Schätzgleichung lautet:

$$\hat{\gamma}^0_{ij} = \arg\min_{\gamma^0_{ij}} \left(\hat{\gamma}_{ij} - \gamma^0_{ij}(K)\right)' \left[V\left(\hat{\gamma}_{ij}\right)\right]^{-1} \left(\hat{\gamma}_{ij} - \gamma^0_{ij}(K)\right) \tag{6.12}$$

$V\left(\hat{\gamma}_{ij}\right)$ ist hierbei die Varianz-Kovarianz-Matrix der nicht restringierten $\hat{\gamma}_{ij}$. Anhand Gleichung (6.12) werden die $\hat{\gamma}^0_{ij}$ generiert, indem die quadrierten gewichteten Differenzen zwischen $\hat{\gamma}_{ij}$ und γ^0_{ij} minimiert werden. Der Ausdruck in der mittleren Klammer ist dabei als Gewichtungsfaktor zu interpretieren, der sicherstellt, dass die Koeffizienten, die geringere Standardfehler haben, stärker gewichtet werden als die Koeffizienten mit größeren Standardfehlern.

6.7 Ermittlung der tatsächlichen Zinsdifferenz

Die Minimum-Distance Schätzung führt zu Koeffizienten, die die Konkavitätsbedingung erfüllen. Nun läßt sich nach Gleichung (6.4) der Optimalzinssatz $\hat{i}$ für jeden landwirtschaftlichen Betrieb ermitteln. Im Anschluss erfolgt dann die Berechnung der Zinsdifferenzen nach Gleichung (6.5). Ein Nachteil der Minimum-Distance Schätzung ist, dass bei dieser Methode nicht die Standardfehler berechnet werden können.

Die Verteilung der Zinsdifferenz zwischen dem $\hat{i}$ und tatsächlichem realem Zinssatz i zeigt eine Tendenz zur Unterkapitalisierung, da der größte Teil der Beobachtungen im positiven Bereich liegt und somit der Schattenzinssatz des Kapitals über dem jeweilig gezahlten Realzinssatz liegt. Dies ist ein erstes Indiz eines Nachfrageüberschusses bei beschränktem Zugang zum Agrarkreditmarkt bzw. für erhöhte Risikoaversion.

Aus diesem Ergebnis lässt sich jedoch noch nicht schließen, dass die Transaktionskosten die Ursache für einen beschränkten Marktzugang der Landwirte sind, da noch die Risikoaversion der Landwirte berücksichtigt werden muss, die eine systematische positive Abweichung des Optimalzinssatzes von dem zu zahlenden Realzinssatz ergibt. Da die Risikoaversion der Landwirte nicht bekannt ist, versuchen wir sie mit Hilfe eines Risikoaufschlages zu simulieren. Auf der Grundlage dieser Simulation lässt sich eine durchschnittliche Verteilung der Differenz zwischen $\hat{i}$ und risikokorrigierten Zinsen unter der Annahme, dass keine Transaktionskosten oder asymmetrische Marktzugangsbarrieren bestehen, ableiten. Vergleicht man dann die tatsächliche empirische Verteilung mit der aus der Monte Carlo Simulation abgeleiteten Verteilung, so lassen sich schließlich aus diesem Vergleich systematische Rückschlüsse bezüglich der o.g. Transaktionskosten und Marktzugangsbarrieren ziehen.

Die Kalkulation von Risikoaufschlägen wird in der Literatur kritisch gesehen: „Nachteilig ist [...] [bei der Kalkulation von Risikoauf- oder abschlägen], dass es keine sinnvoll begründbare Größe für die [Auf- oder] Abschläge gibt und — wichtiger — es nicht voraussehbar ist, welche Sicherheit mit dieser Auswahlregel erzeugt wird“ (Hanf, 1991, S. 41f). Wir wollen jedoch mit Hilfe der Simulation verschiedener Szenarien für Risikoaufschläge testen, ob selbst bei hoher Risikoaversion, d.h. hohen Risikoaufschlägen und einer großen positiven Zinsdifferenz zwischen Optimalzinssatz und gezahltem Realzinssatz, die Höhe der Transaktionskosten auf dem ländlichen Kreditmarkt einen weitaus größeren Einfluss hat und als die entscheidende Determinante der Kapitalausstattung der landwirtschaftlichen Unternehmen identifiziert werden kann.

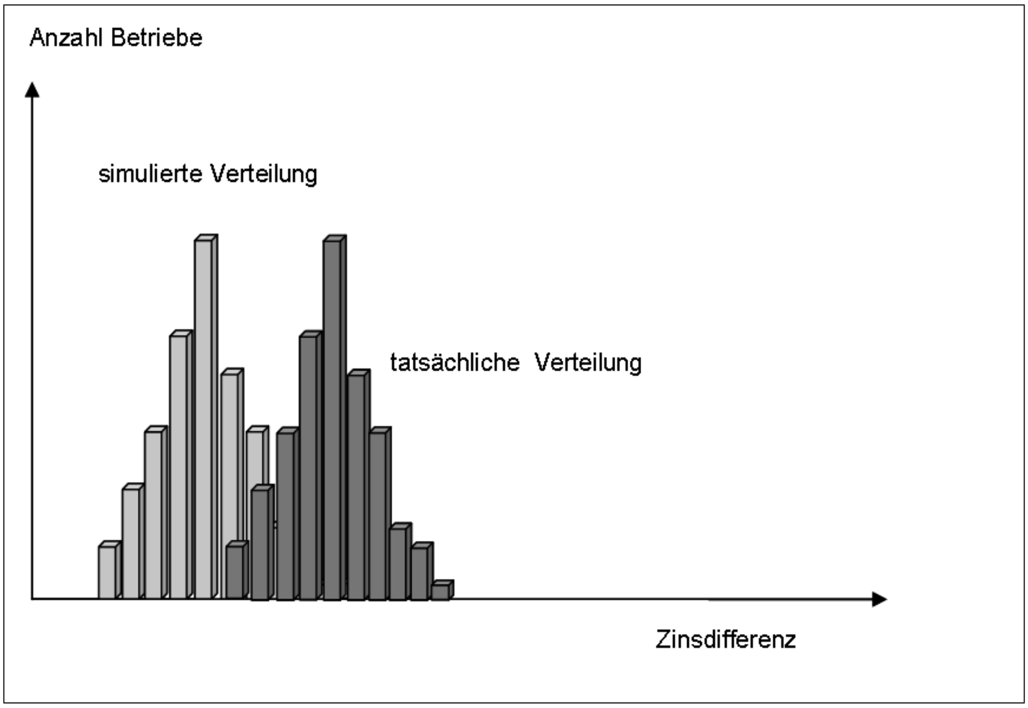
6.8 Monte Carlo Simulation

Mit Hilfe einer Monte Carlo Simulation lässt sich eine hypothetische Verteilung der Zinsdifferenz auf dem polnischen Agrarkreditmarkt nach Gleichung (6.8) generieren. Hierbei wurde angenommen, dass der Störterm ε normalverteilt ist mit $E[\varepsilon] = 0$ und $\sigma_\varepsilon^2 = 0,025^2$ und sowohl positive als auch negative Werte annehmen kann. Für den Risikoaufschlag λ wurden verschiedene Szenarien durchgespielt.

6.9 Vergleich der tatsächlichen mit der simulierten Verteilung

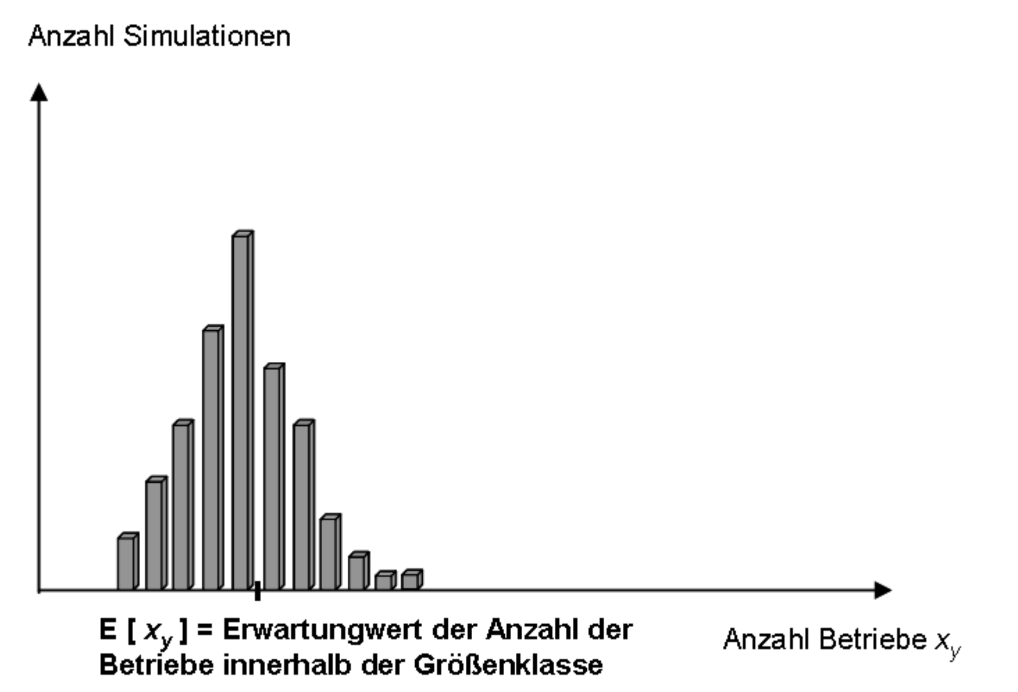
Um die tatsächliche Verteilung der Zinsdifferenz zwischen Optimalzinssatz $\hat{i}$ und Realzinssatz i mit der durchschnittlichen simulierten Verteilung vergleichen zu können, werden die beiden Verteilungen zunächst in Größenklassen eingeteilt, wobei eine Größenklasse y jeweils zwei Prozentpunkte der Zinsdifferenz abdeckt (siehe Abbildung 6.1). Danach wird für jede

Abbildung 6.1: Beispiel für eine tatsächliche und simulierte Verteilung der Zinsdifferenz



Größenklasse y die Anzahl der Simulationen berechnet, bei der x_y Betriebe in der Größenklasse y sind, und der Erwartungswert $E[x_y]$ für die Anzahl der Betriebe x_y in jeder Größenklasse y bestimmt (siehe Abbildung 6.2). Schließlich wird der quadrierte Abstand zwischen diesem Erwartungswert und den einzelnen Simulationen für jede Größenklasse y ermittelt und über die Größenklassen für jede Simulation aufsummiert und die Grenze des 95%-Konfidenzintervalls für die Verteilung dieses Abstandes ermittelt. Zuletzt wird der quadrierte Abstand zwischen dem Erwartungswert und den Werten der tatsächlichen Verteilung analog ermittelt und überprüft, ob der jeweilige Abstand in Grenzen des 95%-Konfidenzintervalls liegt. Liegt der Abstand der tatsächlichen Verteilung innerhalb des 95%-Konfidenzintervalls der simulierten Verteilung der Zinsdifferenz, so kann davon ausgegangen werden, dass die simulierte

Abbildung 6.2: Beispiel für die Verteilung der Simulationen einer bestimmten Größenklasse y



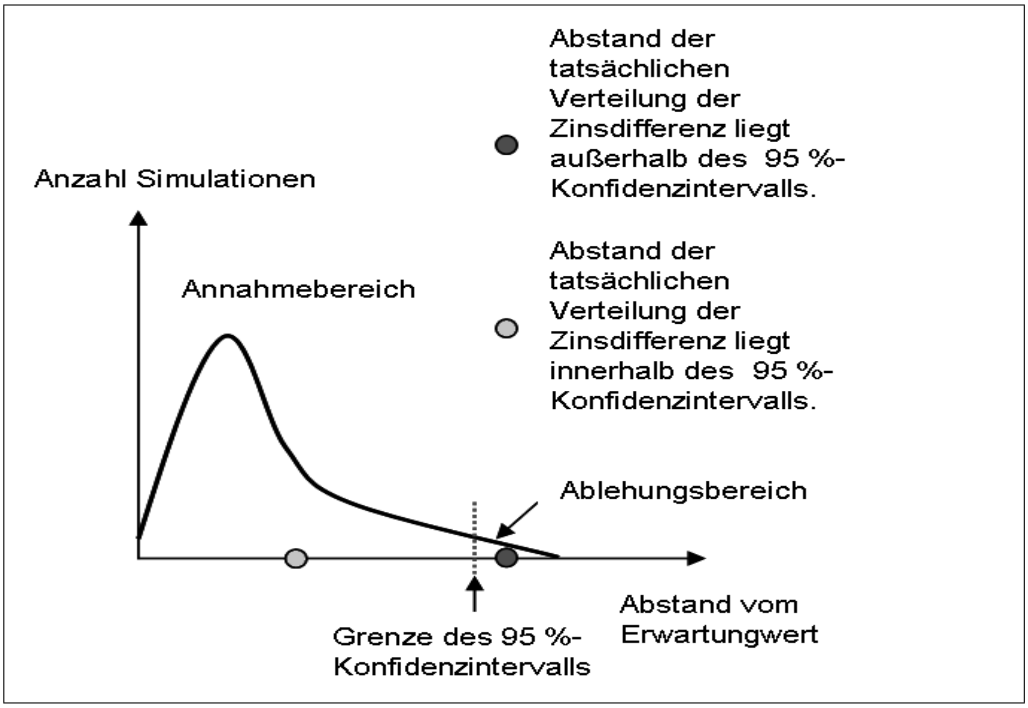
Verteilung der Zinsdifferenz mit der tatsächlichen Zinsdifferenz übereinstimmt, anderenfalls wird von keiner Übereinstimmung ausgegangen (siehe Abbildung 6.3).

6.10 Ergebnisse

Bei den Schätzungen wurde jeweils davon ausgegangen, dass der Risikoaufschlag und die Störgröße normalverteilt sind und der Variationskoeffizient des Risikoaufschlages 0,1 beträgt. Um den Störterm zu simulieren, haben wir eine Standardabweichung von 0,025 angenommen.

Für die polnischen Betriebe konnten wir feststellen, dass die Anpassung der simulierten Verteilung an die tatsächliche Verteilung mit steigendem λ immer besser wurde. Da die Anpassungsschritte mit steigendem λ immer kleiner wurden, strebte der Risikoaufschlag λ gegen unendlich. Dies bedeutet nach unserem Modellansatz, dass selbst wenn wir für die polnischen Landwirte eine hohe Risikoaversion unterstellen würden, der Transaktionskosteneinfluss auf dem ländlichen Kreditmarkt einen weitaus größeren Einfluss hat und als die entscheidende Determinante der Kapitalausstattung der landwirtschaftlichen Unternehmen zu identifizieren ist.

Abbildung 6.3: Prüfung, ob tatsächliche Verteilung der simulierten Verteilung entspricht



Für die schleswig-holsteinischen Testbetriebe konnten wir hingegen einen Risikoaufschlag von $\lambda = 3,0$ ermitteln.[6]

6.11 Schlussbetrachtung

Ziel unseres Beitrages war es, eine Methode zu entwickeln, mit der man den Einfluss von Transaktionskosten auf den ländlichen Kreditmarkt in Polen messen kann. Diese Methode wurde auf polnische Betriebsdaten und, um einen Vergleich zu einem Kreditmarkt zu haben, bei dem das institutionelle Umfeld verhältnismäßig gut entwickelt ist und die Transaktionskosten somit als verhältnismäßig niedrig einzustufen sind, auf Daten der schleswig-holsteinischen Testbetriebsstatistik angewandt. Bei der Simulation des Risikoaufschlages λ sind

[6]Um die Ergebnisse der Analyse der polnischen Betriebsdaten einordnen zu können, wurde die gleiche Schätzmethode auch auf Betriebsdaten des schleswig-holsteinischen Testbetriebsnetzes angewandt. Hierzu standen uns Daten von 721 Testbetrieben aus dem Wirtschaftsjahr 1999/2000 zur Verfügung. Bei der Erzwingung der Konkavität der Produktionsfunktion wurde ebenso wie bei den polnischen Betriebsdaten keine Konvergenz erzielt, so dass ebenfalls wieder die Minimum-Distance Schätzung zum Einsatz kam. Die Schätzung des unrestringierten Modells ist für die schleswig-holsteinischen Betriebe wesentlich besser als für die polnischen Betriebe, fast alle Schätzer sind statistisch signifikant und der R^2-Wert liegt bei 0,88. Auch der Standardfehler für das Wertgrenzprodukt des Kapitals ist mit Werten zwischen 0,01 und 0,05 wesentlich geringer als für die polnische Schätzung.

wir dabei von einem multiplikativen Aufschlag auf den Realzinssatz i ausgegangen. Denkbar wären jedoch auch andere Formen der Berücksichtigung des Risikos, die u.U. unseren Modellansatz noch erweitern und optimieren könnten.

Unsere Analysen bestätigen die Hypothese, dass die Hauptursache für die geringe Investitionstätigkeit in der polnischen Landwirtschaft in den auf dem Agrarkreditmarkt vorherrschenden hohen Transaktionskosten zu suchen ist. Ziel einer effizienten Agrarkreditpolitik sollte es deshalb sein, das institutionelle Umfeld des Agrarkreditmarktes zu verbessern, um die Transaktionskosten auf dem Agrarkreditmarkt zu senken.

Anhang

Abbildung 6.A1: tatsächliche und simulierte Verteilung der Zinsdifferenz für die polnischen Betriebe

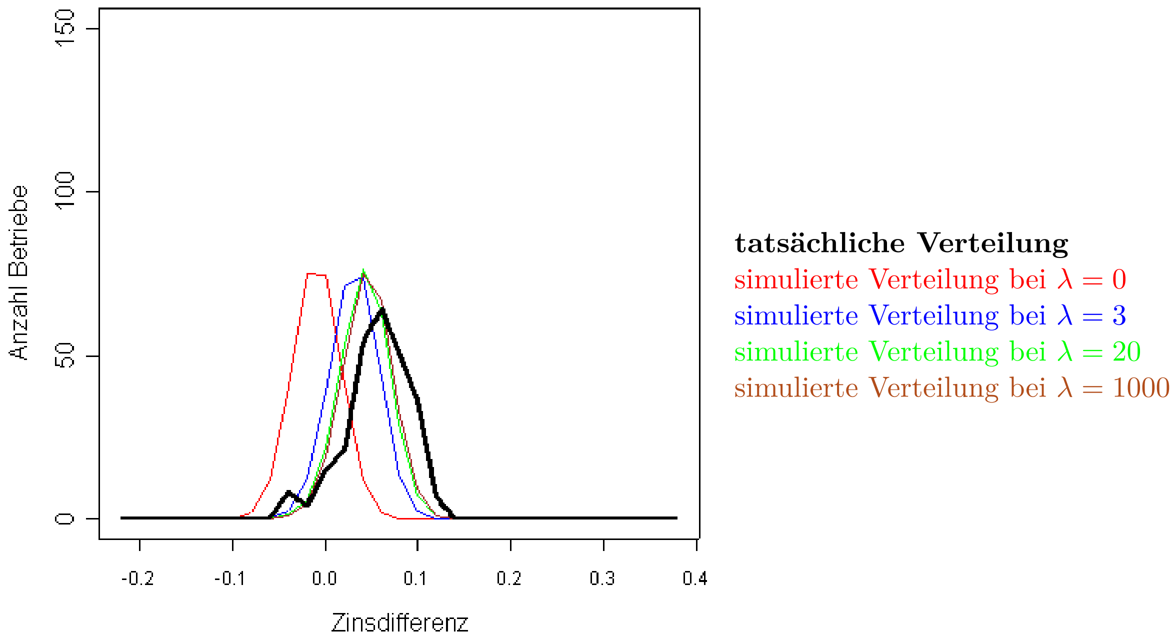
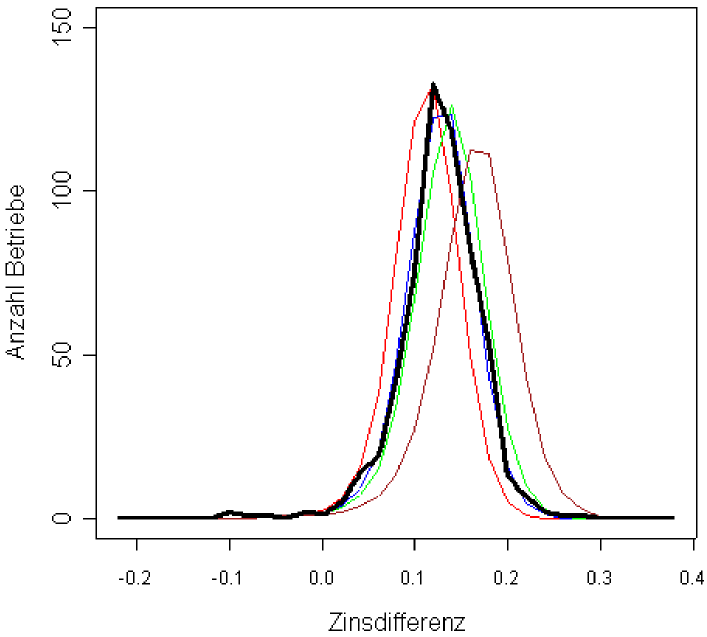


Abbildung 6.A2: tatsächliche und simulierte Verteilung der Zinsdifferenz für die schleswig-holsteinischen Betriebe



Danksagung

Die Autoren möchten dem Institut für Agrarentwicklung in Mittel- und Osteuropa (IAMO) für die Zurverfügungstellung der Datenbasis für diesen Beitrag danken!

Literaturverzeichnis

Brooks, C. (2002): Introductory Econometrics for Finance. Cambridge University Press, Cambridge.

Cooper, J. C. B. (1990): LDC Secondary Market Debt Prices: A Principal Component Analysis. Discussion Paper 10, Department of Economics, Glasgow College, Glasgow.

Fuss, M. und D. McFadden (1978): Production Economics: A Dual Approach to Theory and Applications, Band I: The Theory of Production. North-Holland, Amsterdam.

Hanf, C.-H. (1991): Entscheidungslehre: Einführung in die Informationsbeschaffung, Planung und Entscheidung unter Unsicherheit. 2 Auflage. Oldenbourg, München.

Instytut Ekonomiki Rolnictwa i Gospodarki Zywnosciowej (IERiGZ) (1998): Raporty Rynkowe, Rynek Srodków Produkcji i Uslug dla Rolnictwa - Stan i Perspektywy.

Koebel, B. (1998): Tests of Representative Firm Models: Results for German Manufacturing Industries. In: Journal of Productivity Analysis, 10:251–270.

Koebel, B., M. Falk und F. Laisney (2003): Imposing and Testing Curvature Conditions on a Box-Cox Cost Function. In: Journal of Business and Economic Statistics, 21(2):319–335.

Koester, U. (2001): Agricultural Finance and Institutional Reforms in Transition Economies: The 1990s and Challenges Ahead. In: Quarterly Journal of International Agriculture, 40(4):301–323.

Moosburger, A., A. Balmann und M. Odening (1999): Struktur- und Effizienzwirkungen der einzelbetrieblichen Investitionsförderung in Polen. In: Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., 35:360–365.

Petrick, M. (2000): Are Polish Farmers Credit Constrained? Theoretical Considerations and First Results of a Survey in Western Pomerania (Poland). Paper Presented at the International Scientific Conference 'Regional Structural Transformation of Agriculture and Rural Areas with Regard to Poland's Integration with the EU', Krakow, Poland, 7.–8. 12. 2000.

Puslecki, Z. (2000): Polens Landwirtschaft zwischen Beharrung und EU-Beitritt. In: WeltTrends - Zeitschrift für internationale Politik und vergleichende Studien, 27:95–107.

Kapitel 7

Modeling Farm Households' Price Responses in the Presence of Transaction Costs and Heterogeneity in Labor Markets

Christian H.C.A. Henning und Arne Henningsen

zur Veröffentlichung akzeptiert in: American Journal of Agricultural Economics

Abstract

We develop a farm household model to analyze price responses of farm households. This model incorporates various types of transaction costs as well as labor heterogeneity. Non-proportional variable transaction costs or labor heterogeneity imply that production and consumption decisions become non-separable, even when the household buys or sells labor. An empirical model is estimated using data from Mid-West Poland. The results show that non-proportional variable transaction costs and labor heterogeneity significantly influence household behavior. Not all price elasticities, however, change significantly if these are neglected.

Key words: farm household model, market imperfection, rural labor markets, selectivity, transaction costs

7.1 Introduction

The agricultural development literature has long recognized that rural markets are often underdeveloped or absent. These market imperfections create transaction costs and, if transaction costs are sufficiently high, households find it unprofitable to either buy or sell a good in the market, i.e. remain autarkic (de Janvry, Fafchamps, and Sadoulet, 1991). In this case, production and consumption decisions are no longer separable and conventional microeconomic theory is no longer suitable to model farm household behavior. As a result, farm household models (FHMs) have been developed that explicitly incorporate the interdependency of production and consumption decisions.

Early FHM studies use non-separable FHMs to explain sometimes paradoxical — and even perverse — microeconomic responses of peasants to changes in relative prices (Strauss, 1986; Lopez, 1984; de Janvry, Fafchamps, and Sadoulet, 1991; de Janvry et al., 1992). Several theoretical and empirical studies have used the FHM approach to analyze farm household responses under imperfect labor (Lopez, 1986; Thijssen, 1988; Benjamin, 1992; Jacoby, 1993; Sadoulet, de Janvry, and Benjamin, 1998), capital (de Janvry et al., 1992), or food markets (de Janvry, Fafchamps, and Sadoulet, 1991; Goetz, 1992; Omamo, 1998; Skoufias, 1994; Abdulai and Delgado, 1999). However, non-separability makes theoretical and, in particular, empirical analyses more difficult. Therefore, most empirical analyses assume separable FHMs or use reduced forms of a non-separable FHM.

In contrast to early FHM work, recent studies emphasize transaction costs and institutions in determining households' decisions on market participation (Goetz, 1992; Key, Sadoulet, and de Janvry, 2000; Vakis, Sadoulet, and de Janvry, 2003; Vance and Geoghegan, 2004; Carter and Yao, 2002; Carter and Olinto, 2003). For instance, Key, Sadoulet, and de Janvry (2000) develop a model of supply response when transaction costs cause some producers to buy, others sell, and others do not participate in markets (Key, Sadoulet, and de Janvry, 2000, p. 245). They consider fixed transaction costs (FTC) and proportional transaction costs (PTC) only. Fixed transaction costs are invariant to the quantity of the good traded, whereas proportional transaction costs increase proportionally in quantity. Thus, PTC correspond to constant marginal transaction costs.

An aspect that is conceivable, but has not yet received attention in the FHM literature is the role of non-proportional variable transaction costs (NTC) on production and consumption decisions or market participation. We fill this gap by examining how NTC affect farm household decisions.

We also show that not only transaction costs, which are partly implied by unobserved heterogeneity, but also observed heterogeneity of labor can result in a non-separable FHM. To this end, we construct an FHM, taking into account labor market imperfections via FTC, PTC, NTC, and observed labor heterogeneity. Based on this generalized FHM approach, we derive the following theoretical results: (i) non-separability of production and consumption decisions can occur even if households participate in markets, (ii) imperfect labor markets take a middle ground, with respect to price responses, between standard non-separable FHMs assuming absent labor markets and standard separable FHMs assuming perfect labor markets, and (iii) a test of the joint significance of NTC and heterogeneity for farm household's behavior is possible.

We estimate our generalized FHM approach econometrically using farm household data from Poland. The estimation procedure utilized allows us to consider both potential selectivity and endogeneity problems.

Furthermore, we explicitly test for the significance of NTC and heterogeneity in rural labor markets as well as for the differences between price elasticities calculated for different degrees of labor market imperfection.

7.2 Theoretical Model

In this section we construct a static model of the price responses of farm households in imperfect and perfect labor markets (see also Glauben, Henning, and Henningsen, 2003). To concentrate on the role of labor market constraints, our model ignores some aspects of farmers' decisions, notably (price) risk (Finkelshtain and Chalfant, 1991; Fafchamps, 1992) and credit constraints (Chambers and Lopez, 1987). The farm household is assumed to maximize utility subject to a technology, time, and budget constraint. Therefore, farm households solve the following maximization problem:

$$\max_{\boldsymbol{x},\boldsymbol{c}} U(\boldsymbol{c}) \tag{7.1}$$

subject to

$$G(\boldsymbol{x}, \boldsymbol{r}) = 0 \qquad \text{(production function)} \tag{7.2}$$

$$T_L - |X_L| + X_L^h - X_L^s - C_L \geq 0 \qquad \text{(time constraint)} \tag{7.3}$$

$$P_m C_m \leq P_c X_c + P_a(X_a - C_a) - P_v|X_v| - g(X_L^h) + f(X_L^s) + E \qquad \text{(budget constraint)} \tag{7.4}$$

where $U(\boldsymbol{c})$ is the farm household's utility function, which is monotonically increasing and strictly quasi-concave, and $\boldsymbol{c}$ is a vector of consumption goods consisting of market commodities (C_m), self-produced agricultural goods (C_a), and leisure (C_L).

Production technology is represented by a well-behaved multi-input multi-output production function (7.2) (Lau, 1978a), where $\boldsymbol{x}$ is a vector of production goods, expressed as netputs, and $\boldsymbol{r}$ is a vector of quasi-fixed factors. The farm household produces pure market goods ($X_c > 0$) and goods that are partly consumed by the household ($X_a > 0$). It uses variable intermediate inputs ($X_v < 0$), labor ($X_L < 0$), and the quasi-fixed factors land (R_g) and capital (R_k).

The farm household faces a time constraint (7.3), where T_L denotes total time available. $|X_L| = X_L^f + X_L^h$ is the total of on-farm labor time subdivided into family labor (X_L^f) and hired labor (X_L^h), and X_L^s denotes off-farm family labor. There are four possible regimes of labor market participation. First, the household simultaneously sells family labor and hires labor. Second, farmers neither sell nor hire labor (autarky). Third, households only sell off-farm labor and fourth, they only hire on-farm labor. Earlier studies have neglected the regime in which households simultaneously hire and supply labor. For instance, Sadoulet, de Janvry, and Benjamin (1998) argue that this labor market regime is rarely observed and that their theoretical model cannot explain this specific labor strategy. However, in our data set this regime is rather frequent, with 29% of households falling into that category (table 7.1).[1]

[1] Simultaneously hiring on-farm labor and supplying off-farm labor can be rational with a strictly convex labor cost and a strictly concave labor income function. For instance, if the skills of the household members to work off-farm are very heterogeneous, it is rational to simultaneously supply high-priced labor of well-educated household members and hire cheap agricultural labor (see also Sadoulet, de Janvry, and Benjamin, 1996). A more detailed explanation is provided in Henning and Henningsen (forthcoming).

The budget constraint (7.4) states that a household's consumption expenditures (left-hand side) must not exceed its monetary income (right-hand side). The household may receive income from farming and off-farm employment. In addition, it receives ($E > 0$) or pays ($E < 0$) transfers, which are determined exogenously. Here, P_i, $i \in m, a, c, v$, denote the exogenous consumer and producer prices.

Rural labor markets are often plagued by incomplete formal institutions, which implies transaction costs (Benjamin, 1992; Sadoulet, de Janvry, and Benjamin, 1998; Key, Sadoulet, and de Janvry, 2000). Transaction costs are normally considered as FTC and PTC in existing studies (Key, Sadoulet, and de Janvry, 2000; Vakis, Sadoulet, and de Janvry, 2003). In particular, PTC correspond to transportation and marketing costs, while search, information, negotiation, and bargaining costs as well as screening, enforcement, and supervision costs are generally considered as FTC (Key, Sadoulet, and de Janvry, 2000). Although the concept of FTC and PTC appears intuitive, there is apparently no theoretical justification for excluding NTC ex ante. Empirically there might be some transaction costs that vary non-proportionally with the quantity traded, implying NTC for both on-farm labor demand and off-farm labor supply. Theoretically, it is unclear how the marginal costs vary, i.e. if they are increasing, decreasing, or constant.[2] In this article we present a theoretical framework that considers the impact of NTC on both on-farm labor demand and off-farm labor supply, and also provide an empirical test of their significance.

To formally include NTC as well as FTC and PTC in our model, we denote total variable transaction costs (PTC + NTC) of off-farm employment by $TC_v^s(X_L^s, \boldsymbol{z}_v^s)$ and total variable transaction costs of on-farm labor demand by $TC_v^h(X_L^h, \boldsymbol{z}_v^h)$, where $\boldsymbol{z}_v^s$ and $\boldsymbol{z}_v^h$ denote factors explaining variable transaction costs of the farm household for selling and buying labor, respectively (see Key, Sadoulet, and de Janvry, 2000). For the special case of only PTC these functions are linear in X_L^s and X_L^h, respectively.

Transaction costs are partly implied by unobserved heterogeneity of labor (Spence, 1976; Eswaran and Kotwal, 1986; Frisvold, 1994; Sadoulet, de Janvry, and Benjamin, 1998). However, heterogeneity of labor quality might also have an impact, although it can be observed by employers. For example, family members might have heterogeneous skills to work off-farm, which are generally observable by firms. In such cases, family members would receive different off-farm wage rates corresponding to their observable skills.

If we further assume that family labor is homogeneous regarding farm work, profit maximization implies that the order in which family members work off-farm corresponds to their

[2]We do not intend to provide a comprehensive theory of rural labor market organization and its impact on transaction costs, but rather leave it to future research. Some intuitive examples of NTC are however provided in Henning and Henningsen (forthcoming).

skill levels, further implying that marginal off-farm wage is a step-wise decreasing function of off-farm labor supply. We approximate the step-wise labor wage function by a continuous function. Subtracting marginal transaction costs, we obtain the following effective marginal labor wage function:

$$P_L^s = \overline{P}_L + b^s\left(X_L^s, \boldsymbol{z}_L^s\right) - \frac{\partial TC_v^s\left(X_L^s, \boldsymbol{z}_v^s\right)}{\partial X_L^s}, \tag{7.5}$$

where $\overline{P}_L$ denotes the average regional labor wage, $\boldsymbol{z}_L^s$ denotes the factors explaining heterogeneity of the quality of family labor regarding off-farm work, and $b^s(X_L^s, \boldsymbol{z}_L^s)$ denotes the upward or downward shift of the average labor wage observed by the farm household. We expect that b^s is non-increasing in labor supply, according to our expositions above.

Taking observable heterogeneity and transaction costs into account, the effective revenues from off-farm employment are a function of supplied labor time:

$$f\left(X_L^s, \boldsymbol{z}_L^s, \boldsymbol{z}_v^s, \boldsymbol{z}_f^s\right) = \overline{P}_L X_L^s + \int_0^{X_L^s} b\left(\mathcal{X}, \boldsymbol{z}_L^s\right) \mathrm{d}\,\mathcal{X} - TC_v^s\left(X_L^s, \boldsymbol{z}_v^s\right) - Y^s TC_f^s\left(\boldsymbol{z}_f^s\right), \tag{7.6}$$

where Y^s equals one if $X_L^s > 0$, and zero otherwise; $TC_f^s(\boldsymbol{z}_f^s)$ denotes fixed transaction costs, and $\boldsymbol{z}_f^s$ are factors explaining fixed transaction costs of supplying off-farm labor.

Moreover, observed heterogeneity of on-farm labor might also affect labor demand. For example, some studies (Benjamin, 1992; Deolalikar and Vijverberg, 1983, 1987; Frisvold, 1994) point out that different productivity might be observed for hired and family farm labor. We assume that farm-specific productivity also varies across hired workers. As long as this is unobservable by the farm household, heterogeneity implies transaction costs. However, even if farm households observe farm-specific labor productivity of various workers, it still might affect farmers' price responses if it is not fully reflected in the wage rate. Assuming a constant market wage rate for labor, it is rational to hire workers in the order that corresponds to their on-farm productivity. Under this assumption the marginal cost of an effective unit of on-farm labor is a step-wise increasing function of hired on-farm labor. Again, we approximate this step-wise labor cost function by a continuous function and add marginal transaction costs to obtain the effective marginal wage rate:

$$P_L^h = \overline{P}_L + b^h\left(X_L^h, \boldsymbol{z}_L^h\right) + \frac{\partial TC_v^h\left(X_L^h, \boldsymbol{z}_v^h\right)}{\partial X_L^h}, \tag{7.7}$$

where $\boldsymbol{z}_L^h$ denotes the factors explaining heterogeneity of hired on-farm labor, and $b^h(X_L^h, \boldsymbol{z}_L^h)$ denotes the upward or downward shift of the average regional labor wage observed by the farm household. Again, according to our above expositions, we expect that b^h is non-decreasing in labor demand. Taking into account heterogeneity and variable transaction costs on the labor demand side, the effective labor costs result as a function of demanded labor time:

$$g\left(X_L^h, \boldsymbol{z}_L^h, \boldsymbol{z}_v^h, \boldsymbol{z}_f^h\right) = \overline{P}_L X_L^h + \int_0^{X_L^h} b^h\left(\mathcal{X}, \boldsymbol{z}_L^h\right) \mathrm{d}\,\mathcal{X} + TC_v^h\left(X_L^h, \boldsymbol{z}_v^h\right) + Y^h TC_f^h\left(\boldsymbol{z}_f^h\right), \tag{7.8}$$

where Y^h equals one, if $X_L^h > 0$ and zero otherwise, $TC_f^h(\boldsymbol{z}_f^h)$ denotes fixed transaction costs, and $\boldsymbol{z}_f^h$ are factors explaining fixed transaction costs of demanding on-farm labor.

The higher the NTC or heterogeneity, the higher the decrease in the decision price of off-farm labor induced by increasing labor supply and the higher the increase in the decision price of hired on-farm labor induced by an increasing labor demand. Since it holds $\partial P_L^s / \partial X_L^s = \partial^2 f / \partial X_L^{s\,2}$ and $\partial P_L^h / \partial X_L^h = \partial^2 h / \partial X_L^{h\,2}$, the degree of this market imperfection can be measured by the second-order differentials of f and g. With no heterogeneity and no NTC, both functions are linear and both second-order differentials become zero. Hence, in this case, once households participate in labor markets, marginal off-farm income or marginal costs for hired labor are equal to the exogenously given regional wage rate ($\overline{P}_L$), corrected for proportional transaction costs as well as for household-specific wage shifters. Thus, if households participate in one of the labor markets, the farm household model becomes separable and delivers standard microeconomic comparative static results (Sadoulet, de Janvry, and Benjamin, 1998). Of course, if fixed or proportional transaction costs are too high, households may still abstain from the labor market and stay autarkic, implying a non-separable FHM (Key, Sadoulet, and de Janvry, 2000).

In contrast, when labor markets are imperfectly competitive due to heterogeneity or NTC, both functions are non-linear. In this case, the shadow price of labor (P_L^*) is endogenously determined and production and consumption decisions are determined by solving the utility maximization problem (7.1) to (7.4). Hence, non-separability occurs, even when households participate in labor markets. However, although non-linearity of the f or g function clearly indicates labor market imperfection due to heterogeneity or NTC or both, it is generally impossible to separate the partial impacts of NTC or heterogeneity from observed curvature properties (second-order differentials) of the f and g functions alone.

Theoretically, the curvature properties of the labor revenue function f and the labor cost function g are ambiguous. However, for analytical convenience, we assume f to be concave and g to be convex, since a non-concave labor revenue or a non-convex cost function makes the FHM approach less tractable. Since FTC create discontinuities in the f and g functions, solutions to the maximization problem (7.1) to (7.4) cannot be found by simply solving the first-order conditions. Thus, we follow Key, Sadoulet, and de Janvry (2000) and decompose the solution in two steps. First, we solve for the optimal solution conditional on the labor market participation regime, and then choose the regime that leads to the highest utility. Assuming an interior solution for a given labor market regime (Y^h and Y^s), the optimal quantities of consumption and production goods and the allocation of time are determined by conditions (7.2) to (7.4) and the following equations with $\lambda, \phi, \mu > 0$; $C_m, C_a, C_L, X_c, X_a > 0$; $X_L, X_v < 0$; $X_L^s > 0$ if $Y^s = 1$ and $X_L^s = 0$ otherwise; $X_L^h > 0$ if $Y^h = 1$ and $X_L^h = 0$ otherwise.

$$\frac{\partial U(.)}{\partial C_i} - \lambda P_i = 0 \quad i \in \{m, a, L\} \tag{7.9}$$

$$\phi \frac{\partial G(.)}{\partial X_i} + \lambda P_i = 0 \quad i \in \{c, a, v, L\} \tag{7.10}$$

$$\frac{\partial f(.)}{\partial X_L^s} - P_L^* = 0 \quad \text{if} \quad Y^s = 1 \tag{7.11}$$

$$\frac{\partial g(.)}{\partial X_L^h} - P_L^* = 0 \quad \text{if} \quad Y^h = 1 \tag{7.12}$$

where λ, ϕ are Lagrangian multipliers associated with the budget and the technology constraints, respectively. $P_L^* = \mu/\lambda$ denotes the unobservable shadow wage in the case of non-separability, where μ is the Lagrangian multiplier associated with the time constraint. In the separable version, P_L^* corresponds to the exogenous wage rate corrected for PTC and individual wage shifters.

7.3 Comparative Statics

In general, comparative statics are derived from the first-order conditions (7.2) to (7.4) and (7.9) to (7.12) and thus differ for each labor market regime. However, for simplicity we assume that the farm household simultaneously supplies off-farm labor and demands on-farm labor. Following the standard FHM literature (de Janvry, Fafchamps, and Sadoulet, 1991), comparative statics of a non-separable FHM can be decomposed into the following two components:[3]

$$\frac{\mathrm{d}\, Q}{\mathrm{d}\, P_j} = \left.\frac{\partial Q}{\partial P_j}\right|_{P_L^* = \text{const.}} + \frac{\partial Q}{\partial P_L^*} \frac{\mathrm{d}\, P_L^*}{\mathrm{d}\, P_j}; \quad j \in \{c, a, v, m\}; \quad Q \in \{X_{c,a,v,L}, C_{m,a,L}, X_L^{s,h}\} \tag{7.13}$$

[3] Since derivation of the comparative statics of an FHM is quite standard, we omit a detailed derivation here and only present the main equations. For a detailed derivation see for example Strauss (1986), de Janvry, Fafchamps, and Sadoulet (1991), or Henning (1994).

The first term on the right (direct component) represents the supply or demand reactions to changes in the exogenous prices, assuming a constant labor price (P_L^*). The second term (indirect component) represents the adjustments to the changes in the shadow wage rate caused by changes in the same exogenous price.

Assuming separability, farm household's production and consumption adjustments coincide with the direct component of equation (7.13). In this case, a household's net-labor supply is obtained by subtracting farm labor input ($|X_L|$) and leisure (C_L) from its total labor endowment (T_L).

To determine the indirect component of the non-separable model, we derive the shadow price adjustment by applying the implicit function theorem to the time constraint (7.3) (de Janvry, Fafchamps, and Sadoulet, 1991):

$$\frac{\mathrm{d}\, P_L^*}{\mathrm{d}\, P_j} = -\frac{\dfrac{\partial X_L}{\partial P_j} - \dfrac{\partial C_L}{\partial P_j}}{\dfrac{\partial X_L}{\partial P_L^*} + \dfrac{\partial X_L^h}{\partial P_L^*} - \dfrac{\partial X_L^s}{\partial P_L^*} - \dfrac{\partial C_L^H}{\partial P_L^*}} \tag{7.14}$$

The numerator on the right represents the change in time allocation due to increasing exogenous prices. The denominator of equation (7.14) indicates the change in time allocation caused by changes of the shadow wage rate. Equation (7.14) differs from a corresponding standard non-separable FHM assuming absent labor markets by the term $\Lambda = \partial X_L^h/\partial P_L^* - \partial X_L^s/\partial P_L^*$ in the denominator. This term measures the degree of labor market imperfection due to NTC or heterogeneity. Λ is implicitly determined by the first-order conditions (7.11) and (7.12), whereby: $\partial X_L^s/\partial P_L^* = \left(\partial^2 f/\partial {X_L^s}^2\right)^{-1}$ and $\partial X_L^h/\partial P_L^* = \left(\partial^2 g/\partial {X_L^h}^2\right)^{-1}$. Λ is always positive if f is concave and g is convex. As indicated earlier, the degree of labor market imperfection increases with the second-order differentials, $\partial^2 f/\partial {X_L^s}^2$ and $\partial^2 g/\partial {X_L^h}^2$, measured in absolute terms. In the extreme case of infinitely high NTC and labor heterogeneity, Λ approaches zero; hence, comparative statics of the model in (7.13) approximate the comparative statics derived from an autarkic labor market regime. In the opposite extreme case of zero NTC and perfect labor homogeneity, f and g are linear functions and Λ becomes infinity, implying that the induced shadow wage adjustment (7.14) is zero. Thus, the comparative statics of the model (7.13) would be approximating those of a separable FHM.

We derive the complete comparative statics for all exogenous prices based on equations (7.13) and (7.14) (see Henning and Henningsen, forthcoming). It directly follows from our theoretical analysis above that generally comparative static effects differ across labor market regimes, where the differences between the non-separable and the separable FHM increase with the level of market imperfection due to NTC and heterogeneity.

7.4 Empirical Specification

We fully specify a non-separable farm household model that can be econometrically estimated to assess the question if and to what extent market imperfection influences price responses of farm households. Although our FHM approach includes FTC, PTC, and NTC, as well as labor heterogeneity, our empirical analysis focuses on market imperfection due to NTC and heterogeneity. The empirical specification and estimation strategy are presented in this section and the section thereafter. A more comprehensive derivation is given in Henning and Henningsen (forthcoming).

7.4.1 Production Technology

The production technology (7.2) is represented by a multi-input multi-output profit function from the symmetric normalized quadratic (SNQ) form (Diewert and Wales, 1987, 1992; Kohli, 1993). The corresponding netput equations of the four netputs specified in the theoretical model are given by:

$$\begin{aligned} X_{in}\left(\boldsymbol{p}_{pn}, \boldsymbol{r}_n\right) = \alpha_i + w_n^{-1} \sum_{j \in \{c,a,v,L\}} \beta_{ij} P_{jn} - \frac{1}{2} \theta_i w_n^{-2} \sum_{j \in \{c,a,v,L\}} \sum_{k \in \{c,a,v,L\}} \beta_{jk} P_{jn} P_{kn} \\ + \sum_{j \in \{g,k\}} \delta_{ij} R_{jn} + \frac{1}{2} \theta_i \sum_{j \in \{g,k\}} \sum_{k \in \{g,k\}} \gamma_{jk} R_{jn} R_{kn} \quad \forall\, i \in \{c,a,v,L\} \end{aligned} \tag{7.15}$$

where n indicates the observation (household), $\boldsymbol{p}_{pn} = (P_{an}, P_{cn}, P_{vn}, P_{Ln})$ indicates the netput prices, $\boldsymbol{r}_n = (R_{gn}, R_{kn})$ indicates quasi-fixed factors, $w_n = \sum_{i \in \{c,a,v,L\}} \theta_i P_{in}$ is a factor to normalize prices, $\theta_i = \sum_n P_{in} |X_{in}| / \sum_n \sum_{j \in \{c,a,v,L\}} P_{jn} |X_{jn}|$; $i \in \{c,a,v,L\}$ are predetermined weights of the individual netput prices, and α_i, β_{ij}, δ_{ij}, and γ_{ij} are the parameters to be estimated. To identify all β_{ij}, we impose the restrictions $\sum_{j \in \{c,a,v,L\}} \beta_{ij} \overline{P_j} = 0$; $i \in \{c,a,v,L\}$, where $\overline{P_j}$ are the mean prices (Diewert and Wales, 1987, p. 54). Homogeneity in prices is automatically attained by the functional form and symmetry requires $\beta_{ij} = \beta_{ji}$ $\forall\, i,j \in \{c,a,v,L\}$.

7.4.2 Consumption Decisions

The preferences of the farm households (7.1) and the corresponding consumption decisions are specified by an Almost Ideal Demand System (AIDS) (Deaton and Muellbauer, 1980), i.e. expenditure shares of consumer goods result in:

$$W_{in} = \alpha_i + \sum_{j\in\{m,a,L\}} \gamma_{ij} \ln P_{jn} + \beta_i \ln \frac{Y_n}{\wp_n} \quad \forall\, i \in \{m, a, L\} \tag{7.16}$$

$$\text{with } \ln \wp_n = \alpha_0 + \sum_{i\in\{m,a,L\}} \alpha_i \ln P_{in} + \frac{1}{2} \sum_{i\in\{m,a,L\}} \sum_{j\in\{m,a,L\}} \gamma_{ij} \ln P_{in} \ln P_{jn} \tag{7.17}$$

where $W_{in} = P_{in}C_{in}/Y_n$; $i \in \{m, a, L\}$ are the expenditure shares, Y_n indicates full income, $\wp_n$ is the translog consumer price index, P_{in}; $i \in \{m, a, L\}$ indicates the consumer prices, and α_i, β_i, and γ_{ij} are the parameters to be estimated. Adding-up requires $\sum_{i\in\{m,a,L\}} \alpha_i = 1$, $\sum_{i\in\{m,a,L\}} \beta_i = 0$, $\sum_{i\in\{m,a,L\}} \gamma_{ij} = 0$, and homogeneity in prices requires $\sum_{j\in\{m,a,L\}} \gamma_{ij} = 0$, and symmetry requires $\gamma_{ij} = \gamma_{ji}\ \forall\, i, j \in \{m, a, L\}$.

7.4.3 Labor Market Decisions

To allow imperfect labor markets due to FTC, PTC, and NTC, as well as (observed) heterogeneity, we assume a quadratic form for the labor income function f in (7.6) and the labor cost function g in (7.8), which can be interpreted as second-order approximations of the true labor cost and income functions, respectively. According to our theoretical expositions above, assuming quadratic f and g functions implies that the shadow wage functions (7.5) and (7.7) are linear:

$$P_L^* = \beta_0^s + X_L^s \beta_1^s + \boldsymbol{z}^{s\prime} \boldsymbol{\beta}^s \tag{7.18}$$

$$P_L^* = \beta_0^h + X_L^h \beta_1^h + \boldsymbol{z}^{h\prime} \boldsymbol{\beta}^h \tag{7.19}$$

As in equation (7.5), the vector $\boldsymbol{z}^s$ includes factors that explain variable transaction costs (PTC and NTC) of supplying labor ($\boldsymbol{z}_v^s$) and the average skill level of a farm household ($\boldsymbol{z}_L^s$) as well as a proxy for the average regional wage level ($\widetilde{P}_L$). Analogously, as in equation (7.7), the vector $\boldsymbol{z}^h$ includes factors explaining PTC and NTC of hiring labor ($\boldsymbol{z}_v^h$) and the average skill of hired on-farm labor ($\boldsymbol{z}_L^h$) as well as a proxy for the average regional wage level ($\widetilde{P}_L$). Moreover, since the quadratic functions are second-order approximations of the true f and g functions, their (local) curvature properties are fully captured by the coefficients β_1^s and β_1^h, respectively. Accordingly, we can separately test for the significance of NTC and heterogeneity in off-farm and in on-farm labor markets with a t-test. The null hypotheses correspond to $H_0 : \beta_1^h = 0$ and $H_0 : \beta_1^s = 0$.[4] Non-separability is implied if both null hypotheses are rejected. However, even if one of the null hypotheses cannot be rejected, non-separability can still occur if the farm household does not participate in the corresponding labor market owing to high fixed or proportional transaction costs.[5]

7.5 Estimation Strategy

The econometric estimation of the empirical model specified above (7.15–7.19) is not straightforward, since shadow prices of labor cannot be observed directly. Therefore, we use a two-stage estimation strategy. We estimate shadow prices of labor at the first stage and at the second stage we estimate separately the SNQ profit function (7.15), the Almost Ideal Demand System (7.16, 7.17) and the linear labor wage equations (7.18, 7.19).

7.5.1 Estimating Shadow Values of Labor (Stage 1)

We follow Lopez (1984) to estimate the shadow prices of labor and estimate a restricted profit function with labor as a quasi-fixed input. Assuming constant returns to labor, Lopez (1984) derived the shadow wages of the households as shadow price of labor on the farm. The netput quantities per unit of labor that correspond to an SNQ profit function are

$$\begin{aligned}\frac{X_{in}\left(\boldsymbol{p}_{pn}, \boldsymbol{r}_n, X_{Ln}\right)}{X_{Ln}} &= \alpha_i + w_n^{-1} \sum_{j\in\{c,a,v\}} \beta_{ij} P_{jn} - \frac{1}{2}\theta_i w_n^{-2} \sum_{j\in\{c,a,v\}} \sum_{k\in\{c,a,v\}} \beta_{jk} P_{jn} P_{kn} \\ &\quad + \sum_{j\in\{g,k\}} \delta_{ij} R_{jn} + \frac{1}{2}\theta_i \sum_{j\in\{g,k\}} \sum_{k\in\{g,k\}} \gamma_{jk} R_{jn} R_{kn} \quad \forall\, i \in \{c,a,v\}\end{aligned} \tag{7.20}$$

[4]Our estimation strategy does not permit the estimation of FTC, since TC_f^s and TC_f^h cannot be identified. However, because we are only interested in the impact of imperfect labor markets on price responses, we do not need to identify fixed transaction costs at this stage and we let them be captured by exogenous transfers (E). The simultaneous estimation of FTC, PTC, and NTC is an interesting research topic (see Vakis, Sadoulet, and de Janvry, 2003), which will require more elaboration in future work.

[5]Non-linearity of the labor revenue and labor cost functions is a sufficient, but not a necessary condition for non-separability. It is, however, a necessary condition if households participate in labor markets. Even if the labor revenue and labor cost function are both linear, fixed or proportional transaction costs could be so high that farms abstain from labor markets and thus, their production and consumption decisions are no longer separable. Hence, if our statistical test rejects linearity of the labor revenue and labor cost functions, we can conclude that the FHM is generally non-separable. However, if our test does not reject linearity, we can conclude that the FHM is separable for households that participate in labor markets; nevertheless non-separability could still be observed in autarkic households. Other tests of separability have been suggested for the latter case (see for example Benjamin, 1992). However, we did not apply these additional tests because in our specific empirical application our test was sufficient to identify non-separability (see section "Data and Empirical Results").

where $\boldsymbol{p}_{pn} = (P_{an}, P_{cn}, P_{vn})$ indicates the netput prices. Parameters in equation (7.20) are analogously defined and we impose the analogous restrictions as in equation (7.15).

Finally, the shadow prices of labor can be obtained from the estimation results by

$$P^*_{Ln} = \frac{\partial \widehat{\Pi}_n \left(\boldsymbol{p}_{pn}, \boldsymbol{r}_n, X_{Ln}\right)}{\partial X_{Ln}} \tag{7.21}$$

where $\widehat{\Pi}_n(\boldsymbol{p}_{pn}, \boldsymbol{r}_n, X_{Ln}) = \sum_{i \in \{c,a,v\}} P_{in} \widehat{X}_{in}$ is the fitted variable profit of the n^{th} farm and $\widehat{X}_{in}(\boldsymbol{p}_{pn}, \boldsymbol{r}_n, X_{Ln})$ are the fitted values of the netput quantities.

Microeconomic theory generally requires that profit functions are convex in all netput prices, which is not the case in many empirical estimations. Therefore, we impose convexity of the profit function (7.20), applying a three-step procedure suggested by Koebel, Falk, and Laisney (2003) based on the minimum distance and asymptotic least squares estimation (Gourieroux, Monfort, and Trognon, 1985; Kodde, Palm, and Pfann, 1990).[6] First, the unrestricted (linear) netput equations are estimated to calculate the Hessian matrix of the unrestricted profit function. Second, we minimize the weighted difference between this unrestricted Hessian and a Hessian that is restricted to be positive semi-definite by the Cholesky factorization. Third, restricted coefficients are identified by an asymptotic least squares (ALS) framework. The weighting matrix for the minimization of the difference between the unrestricted and the restricted Hessian matrix is the inverse of the variance-covariance matrix of the Hessian matrix, which can be derived from the coefficient variance-covariance matrix of the unrestricted estimation. The variance covariance matrix of the coefficients is obtained by bootstrapping (Efron, 1979; Efron and Tibshirani, 1993).

7.5.2 Farm Technology (Stage 2a)

Given the estimated shadow prices of labor, we estimate the SNQ netput equations (7.15). Again, we impose convexity with the method of Koebel, Falk, and Laisney (2003). However, the price of labor (P^*_L) is endogenous and a generated regressor. We use a three-stage least squares (3SLS) estimation with the variables $\boldsymbol{z}$ (see below) as instrumental variables for P^*_L,

[6] We first tried to impose convexity by a non-linear estimation using the Cholesky decomposition (Lau, 1978b). However, the estimation of the restricted non-linear netput equations did not converge. The new procedure suggested by Koebel, Falk, and Laisney (2003) circumvents this non-linear estimation and is asymptotically equivalent to a (successful) non-linear estimation with convexity imposed. To retain convexity of the SNQ profit function, it would be sufficient to minimize the difference between the estimated (unrestricted) β coefficients and the (linearly independent) values of a restricted β coefficient matrix (Koebel, 1998). However, this procedure adjusts only the β-coefficients, while the approach of Koebel, Falk, and Laisney (2003) adjusts *all* coefficients. Thus, the fit of the constrained model is much better, due to the flexibility of the other coefficients. Both approaches yield the same βs.

to account for the endogeneity and the generation of P_L^* (Pagan, 1984) and to allow for contemporaneous correlation of the disturbance terms.

7.5.3 Consumption (Stage 2b)

Analogously, given the estimated shadow prices of leisure (labor), we estimate the demand system (7.16, 7.17). In addition to P_L^* being endogenous and a generated regressor, the full income variable (Y) in the consumption decision specification might be endogenous and depends on P_L^*. To avoid estimation biases, we utilize a three-stage least squares (3SLS) estimation, in which we use the variables $\boldsymbol{z}$ (see below) as instruments for P_L^* and Y. To avoid non-linear estimation, the share equations of the AIDS are estimated by the "Iterated Linear Least Squares Estimator" (ILLE) proposed by Blundell and Robin (1999).

7.5.4 Labor Market Decisions (Stage 2c)

Given the estimated shadow prices of labor, we estimate the two linear labor wage functions (7.18) and (7.19). However, these estimations might be plagued by a sample selection bias and an endogeneity problem.[7]

The endogeneity problem arises because the regressors X_L^s and X_L^h are probably correlated with the disturbance terms. To overcome this problem, we use a 2SLS estimation and substitute fitted values $(\widehat{X}_L^s, \widehat{X}_L^h)$ for the observed quantities of supplied and hired labor (X_L^s, X_L^h). According to our theory, the optimal labor market allocation (X_L^s, X_L^h) of households that supply and demand labor simultaneously depends on the first-order conditions (7.9) – (7.12). For households that only supply labor, the optimal amount of supplied labor (X_L^s) depends only on conditions (7.9) – (7.11), while for households that only demand labor, the optimal quantity of hired labor (X_L^h) depends only on conditions (7.9), (7.10), and (7.12). Therefore, the impact of exogenous variables on the amount of traded labor (X_L^s, X_L^h) depends on the labor market regime. Hence, the first stage of this 2SLS estimation corresponds to a switching regression model.

The sample selection bias occurs because these equations can only be estimated for households that participate in labor markets. To correct for selectivity, we apply an extended

[7] The deviations between the estimated and the unobserved (true) shadow prices of labor get a part of the regular error terms ν^s and ν^h of the shadow price equations (7.24) and (7.25). We assume that these deviations are neither correlated with the regressors $\boldsymbol{z}^s$ and $\boldsymbol{z}^h$ nor with the variables used as instruments for X_L^s and X_L^h in the 2SLS estimation ($\boldsymbol{z}_x^b$, $\boldsymbol{z}_x^s$, $\boldsymbol{z}_x^h$). Note that we do not have to assume that the deviations are uncorrelated with the regressors X_L^s and X_L^h because X_L^s and X_L^h are not used as instruments in the 2SLS estimation.

Heckman procedure and add selectivity terms (λ) to these equations, which can be interpreted as an extension of the two-stage probit method for simultaneous equation models with selectivity suggested by Lee, Maddala, and Trost (1980). Assumptions about the error terms are given in Henning and Henningsen (forthcoming). Overall, a consistent estimation of these functions corresponds to the joint estimation of the following eight equations:
Market participation equations (estimated as a bivariate probit model):

$$Y^{s*} = \boldsymbol{z}'\boldsymbol{\gamma}^s + \varepsilon^s \quad \text{with } Y^{s*} > 0 \text{ if } X_L^s > 0 \text{ and } Y^{s*} \leq 0 \text{ if } X_L^s = 0 \tag{7.22}$$

$$Y^{h*} = \boldsymbol{z}'\boldsymbol{\gamma}^h + \varepsilon^h \quad \text{with } Y^{h*} > 0 \text{ if } X_L^h > 0 \text{ and } Y^{h*} \leq 0 \text{ if } X_L^h = 0 \tag{7.23}$$

Shadow wage equations (second stage of the 2SLS estimation):

$$P_L^* = \beta_0^s + \widehat{X}_L^s \beta_1^s + \boldsymbol{z}^{s\prime}\boldsymbol{\beta}^s + \sigma^s \lambda^s + \nu^s \quad \text{if} \quad Y^{s*} > 0 \tag{7.24}$$

$$P_L^* = \beta_0^h + \widehat{X}_L^h \beta_1^h + \boldsymbol{z}^{h\prime}\boldsymbol{\beta}^h + \sigma^h \lambda^h + \nu^h \quad \text{if} \quad Y^{h*} > 0 \tag{7.25}$$

Labor supply and demand equations (first stage of the 2SLS estimation):

$$X_L^s = \boldsymbol{z}_x^{b\prime}\boldsymbol{\delta}_s^b + \sigma_s^{bs}\lambda^{bs} + \sigma_s^{bh}\lambda^{bh} + \xi_s^b \quad \text{if} \quad Y^{s*} > 0 \wedge Y^{h*} > 0 \tag{7.26}$$

$$X_L^s = \boldsymbol{z}_x^{s\prime}\boldsymbol{\delta}_s^s + \sigma_s^{ss}\lambda^{ss} + \sigma_s^{sh}\lambda^{sh} + \xi_s^s \quad \text{if} \quad Y^{s*} > 0 \wedge Y^{h*} \leq 0 \tag{7.27}$$

$$X_L^h = \boldsymbol{z}_x^{b\prime}\boldsymbol{\delta}_h^b + \sigma_h^{bs}\lambda^{bs} + \sigma_h^{bh}\lambda^{bh} + \xi_h^b \quad \text{if} \quad Y^{h*} > 0 \wedge Y^{s*} > 0 \tag{7.28}$$

$$X_L^h = \boldsymbol{z}_x^{h\prime}\boldsymbol{\delta}_h^h + \sigma_h^{hs}\lambda^{hs} + \sigma_h^{hh}\lambda^{hh} + \xi_h^h \quad \text{if} \quad Y^{h*} > 0 \wedge Y^{s*} \leq 0 \tag{7.29}$$

where $\boldsymbol{z} = (1, \boldsymbol{z}^{\pi\prime}, \boldsymbol{z}^{u\prime}, \boldsymbol{z}^{s\prime}, \boldsymbol{z}^{h\prime}, \boldsymbol{z}_f^{s\prime}, \boldsymbol{z}_f^{h\prime})'$ are factors influencing labor market participation; $\boldsymbol{z}_x^b = (1, \boldsymbol{z}^{\pi\prime}, \boldsymbol{z}^{u\prime}, \boldsymbol{z}^{s\prime}, \boldsymbol{z}^{h\prime})'$, $\boldsymbol{z}_x^s = (1, \boldsymbol{z}^{\pi\prime}, \boldsymbol{z}^{u\prime}, \boldsymbol{z}^{s\prime})'$, and $\boldsymbol{z}_x^h = (1, \boldsymbol{z}^{\pi\prime}, \boldsymbol{z}^{u\prime}, \boldsymbol{z}^{h\prime})'$ are factors influencing the quantity of supplied and hired labor (depending on the labor market regime); all ε, ν, and ξ denote the error terms; all γ, β, σ, and δ are parameters or parameter vectors to be estimated, and the selectivity terms are

$$\lambda^s = \frac{\phi\left(\boldsymbol{z}'\boldsymbol{\gamma}^s\right)}{\Phi\left(\boldsymbol{z}'\boldsymbol{\gamma}^s\right)}, \qquad \lambda^h = \frac{\phi\left(\boldsymbol{z}'\boldsymbol{\gamma}^h\right)}{\Phi\left(\boldsymbol{z}'\boldsymbol{\gamma}^h\right)} \tag{7.30}$$

$$\lambda^{bs} = \frac{\phi\left(\boldsymbol{z}'\boldsymbol{\gamma}^s\right)\Phi\left(\frac{\boldsymbol{z}'\boldsymbol{\gamma}^h-\rho\boldsymbol{z}'\boldsymbol{\gamma}^s}{\sqrt{1-\rho^2}}\right)}{\Phi_2\left(\boldsymbol{z}'\boldsymbol{\gamma}^s,\boldsymbol{z}'\boldsymbol{\gamma}^h\right)}, \qquad \lambda^{bh} = \frac{\phi\left(\boldsymbol{z}'\boldsymbol{\gamma}^h\right)\Phi\left(\frac{\boldsymbol{z}'\boldsymbol{\gamma}^s-\rho\boldsymbol{z}'\boldsymbol{\gamma}^h}{\sqrt{1-\rho^2}}\right)}{\Phi_2\left(\boldsymbol{z}'\boldsymbol{\gamma}^s,\boldsymbol{z}'\boldsymbol{\gamma}^h\right)} \tag{7.31}$$

$$\lambda^{ss} = \frac{\phi\left(\boldsymbol{z}'\boldsymbol{\gamma}^s\right)\Phi\left(\frac{-\boldsymbol{z}'\boldsymbol{\gamma}^h+\rho\boldsymbol{z}'\boldsymbol{\gamma}^s}{\sqrt{1-\rho^2}}\right)}{\Phi_2^*\left(\boldsymbol{z}'\boldsymbol{\gamma}^s,-\boldsymbol{z}'\boldsymbol{\gamma}^h\right)}, \qquad \lambda^{sh} = -\frac{\phi\left(\boldsymbol{z}'\boldsymbol{\gamma}^h\right)\Phi\left(\frac{\boldsymbol{z}'\boldsymbol{\gamma}^s-\rho\boldsymbol{z}'\boldsymbol{\gamma}^h}{\sqrt{1-\rho^2}}\right)}{\Phi_2^*\left(\boldsymbol{z}'\boldsymbol{\gamma}^s,-\boldsymbol{z}'\boldsymbol{\gamma}^h\right)} \tag{7.32}$$

$$\lambda^{hs} = -\frac{\phi\left(\boldsymbol{z}'\boldsymbol{\gamma}^s\right)\Phi\left(\frac{\boldsymbol{z}'\boldsymbol{\gamma}^h-\rho\boldsymbol{z}'\boldsymbol{\gamma}^s}{\sqrt{1-\rho^2}}\right)}{\Phi_2^*\left(-\boldsymbol{z}'\boldsymbol{\gamma}^s,\boldsymbol{z}'\boldsymbol{\gamma}^h\right)}, \qquad \lambda^{hh} = \frac{\phi\left(\boldsymbol{z}'\boldsymbol{\gamma}^h\right)\Phi\left(\frac{-\boldsymbol{z}'\boldsymbol{\gamma}^s+\rho\boldsymbol{z}'\boldsymbol{\gamma}^h}{\sqrt{1-\rho^2}}\right)}{\Phi_2^*\left(-\boldsymbol{z}'\boldsymbol{\gamma}^s,\boldsymbol{z}'\boldsymbol{\gamma}^h\right)} \tag{7.33}$$

where $\phi()$ and $\Phi()$ denote the probability density (pdf) and cumulative distribution (cdf) function of a standard normal distribution, respectively, and Φ_2 and Φ_2^* are the cumulative distribution (cdf) functions of a bivariate standard normal distribution with correlations ρ and $-\rho$, respectively. A detailed derivation of the selectivity terms is available in Henning and Henningsen (forthcoming).[8] Equations (7.30) to (7.33) are used to compute the selectivity terms $(\widehat{\lambda})$, which are then substituted for the true λs in equations (7.24) to (7.29). The estimated results of equations (7.26) to (7.29) are then used to obtain fitted values $(\widehat{X}_L^s, \widehat{X}_L^h)$ that are used to estimate the second stage of the 2SLS estimation of equations (7.24) and (7.25). Finally, the variance covariance matrix of the second stage coefficients are computed with the formula given in Lee, Maddala, and Trost (1980) to obtain consistent standard errors.

7.6 Data and Empirical Results

Data are based on an accounting survey of 202 agricultural households in several regions around Poznan (Mid-West Poland) in 1994. The data were collected by the Institute for Agriculture and Food Industries in Warsaw (IERiGZ, 1995). Additional regional data are

[8] We thank Awudu Abdulai, who pointed out that Saha, Love, and Schwart (1994) analyze a similar sample selection problem. In particular, they suggest an extended Heckman procedure, which is also applied by Abdulai, Monnin, and Gerber (2005). Although we have been stimulated by their work, we actually derived slightly different selectivity terms. To compare our results with the results of Saha, Love, and Schwart (1994) we calculated the conditional expectation values by numerical integration and Monte Carlo simulation using the (free) statistical software R (R Development Core Team, 2005, see also http://www.r-project.org), and the add-on packages adapt (Genz et al., 2005), mvtnorm (Genz, Bretz, and Hothorn, 2005), and MASS (Venables and Ripley, 2002). While our formula perfectly fits the numerical calculations, the formula of Saha, Love, and Schwart (1994) did not.

Table 7.1: Characteristics of the Different Labor Regimes

Variable	Unit	All	Sup. & Dem.	Only Sup.	Only Dem.	Autarkic
Number		199	57	47	61	34
N_k	number	1.3	1.5	1.3	1.4	0.7
N_w	number	2.8	2.8	3.2	2.4	3.0
N_o	number	0.7	0.6	0.6	0.8	0.7
A_h	years	43	41	44	43	45
T_L	hours	11399	11110	12891	10082	12185
$\|X_L\|$	hours	3686	3579	3372	4040	3668
X_L^h	hours	211	278	0	430	0
X_L^s	hours	446	515	1266	0	0
X_L^n	hours	235	237	1266	-430	0
X_L^f	hours	3475	3301	3372	3610	3668
C_L	hours	7478	7295	8254	6473	8517
P_mC_m	1000 PLZ	91469	105939	78012	97792	74467
P_aC_a	1000 PLZ	19041	18487	19245	19939	18076
P_cX_c	1000 PLZ	132258	157581	65883	180020	95869
P_aX_a	1000 PLZ	212570	220643	123997	300046	164531
$P_v\|X_v\|$	1000 PLZ	211960	232143	117552	299629	151343
R_g	ha	14.7	16.9	9.4	18.3	11.7
R_k	1000 PLZ	649191	788881	425398	816534	424132
R_k/R_g	1000 PLZ / ha	46921	49666	48516	48134	37938
N_c	number	0.9	1.0	0.8	0.9	0.8
W_u	%	19	20	19	18	20
W_i	km/100 km^2	58	55	60	60	57
W_t	1/1000 popul.	48	47	49	49	47
W_r	%	45	44	50	43	46
$\widetilde{P}_L$	Poland = 100	88	85	90	89	88
P_L^*	1000 PLZ/h	38	46	30	44	28

Note: Calculations are based on IERiGZ (1995). PLZ = Polish Zloty. Variables: N_k = number of family members up to 14 years, N_w = number of family members between 15 and 60 years, N_o = number of family members older than 60 years, A_h = age of the household head, T_L = total time available, $|X_L|$ = labor input on the farm, X_L^h = hired labor, X_L^s = supplied labor, X_L^n = net supplied labor, X_L^f = family labor input on the farm, C_L = leisure, P_mC_m = value of consumed market goods, P_aC_a = value of consumed self-produced goods, P_cX_c = value of produced crop products, P_aX_a = value of produced animal products, $P_v|X_v|$ = value of utilized variable inputs, R_g = amount of land of the farm, R_k = amount of capital of the farm, N_c = number of cars owned by the household, W_u = regional unemployment rate, W_i = regional density of the road and railroad network, W_t = regional density of telephones, W_r = proportion of the population that lives in rural areas, $\widetilde{P}_L$ = relative average regional wage level, $P_\cdot^*$ = endogenous shadow price of labor.

taken from Glowny Urzad Statystyczny (1996) and Zawadzki (1994). Sample characteristics of different labor market regimes are presented in table 7.1.

The empirical specification of the theoretical model is as follows. On the production side, market goods (X_c) consist of all crop products, while animal products are considered as partly home-consumed goods (X_a). All relevant variable inputs of the farms are subsumed in netput X_v. Labor (X_L) includes both family (X_L^f) and hired labor (X_L^h). Land (R_g) and capital (R_k) are considered as quasi-fixed factors. On the consumption side, C_m includes all purchased consumption goods. The self-produced goods (C_a) correspond conceptually to the home-consumed animal products (X_a). The amount of leisure (C_L) is determined by calculating the yearly available time (T_L) of households minus on-farm (X_L^f) and off-farm (X_L^s) family labor.[9]

The variables $\boldsymbol{z}^\pi$ influencing the shadow price of labor from the production side include land and capital endowments (R_g, R_k) as well as variable output and input prices (P_c, P_a, P_v). The variables $\boldsymbol{z}^u$ influencing the shadow price from the consumer side include household composition and consumer prices. In particular, household composition is measured by the number of family members up to 14 years (N_k), between 15 and 60 years (N_w), and older than 60 years (N_o), as well as sex (D_f), age (A_h), and age squared (A_h^2) of the household head, because these variables might influence the preferences for leisure.

The variable and fixed transaction costs on the labor markets ($\boldsymbol{z}_v^s, \boldsymbol{z}_v^h, \boldsymbol{z}_f^s, \boldsymbol{z}_f^h$) are explained by the number of cars owned by the household (N_c), the regional density of the road and railroad network (W_i), the regional number of telephones per 1,000 population (W_t), the regional unemployment rate (W_u), and the proportion of the population that lives in rural areas (W_r). Furthermore, we assume that the average off-farm skill level of farm households ($\boldsymbol{z}_L^s$) depends on the number of family members that are of working age (N_w), the number of family members older than 60 years (N_o), and the average level of human capital. Since no data on education is available, we follow Vakis, Sadoulet, and de Janvry (2003) and interpret sex (D_f), age (A_h), and age squared (A_h^2) of the household head as an indicator of average human capital. Finally, the average skill level of hired workers ($\boldsymbol{z}_L^h$) is explained by the mechanization on the farm, measured as capital intensity (R_k/R_g).

[9] It is assumed that each household member between 15 years and 60 years has 10 hours per day and each household member older than 60 years has five hours per day available for work and/or leisure. The annual available time of the household is calculated by multiplying the total hours per day of all household members by 365. We use the share of off-farm labor in total labor endowment (X_L^s/T_L) instead of the absolute amount of supplied labor (X_L^s) as an explanatory variable in the off-farm labor wage equation to account for different household sizes. Hence, we assume that the share of skilled and unskilled labor in the total household would not significantly vary with the family size. Using the absolute amount of off-farm labor supply instead does not change the main results, i.e. significant and negative impact on the effective off-farm wage rate.

The sample contains two farms that do not produce any animal products, which are removed to provide a more homogeneous sample and to avoid imputing the unknown prices of animal products.

7.6.1 Estimation results

This section presents the main estimation results. More detailed results are available in Henning and Henningsen (forthcoming). All estimations and calculations are carried out by the (free) statistical software "R" (R Development Core Team, 2005, see also http://www.r-project.org), using the add-on packages "micEcon" (Henningsen and Toomet, 2005), "systemfit" (Hamann and Henningsen, 2005), and "VGAM" (Yee and Wild, 1996).

The three netput equations of the SNQ profit function (7.20) are estimated in the first step. The shadow prices of labor calculated from the restricted profit function have reasonable values for all but one farm household. This household has a negative shadow price and is therefore removed from the sample. Hence, the sample used includes 199 farm households.

Tables 7.2 and 7.3 present the estimates of the restricted second-step profit function (7.15) and of the Almost Ideal Demand System (7.16, 7.17), respectively.

Table 7.4 presents the estimates of the off-farm and on-farm labor wage functions. Since the focus of this paper is on market imperfection due to NTC and heterogeneity, the parameters $\beta_1^s = \partial P_L^* / \partial X_L^s$ and $\beta_1^h = \partial P_L^* / \partial X_L^h$ are of particular interest. Recall that these coefficients measure the degree of market imperfection due to NTC and heterogeneity and thus are of particular relevance. The other coefficients, measuring the effects of the variables $\boldsymbol{z}$ on labor market participation decisions and the influence of the variables $\boldsymbol{z}^s$ and $\boldsymbol{z}^h$ on the shadow prices of labor, are only of secondary interest and are explained in Henning and Henningsen (forthcoming).[10]

As can be seen from table 7.4, the effect of labor supply on the off-farm wage rate (β_1^s) is significantly negative. This indicates a concave labor revenue function and, hence, increasing marginal NTC or heterogeneity of off-farm labor skills. If an average household increases off-farm labor supply by 1%, marginal revenue falls by 0.075%. The estimated parameter of the inverse Mill's ratio is not significantly different from zero, implying no sample selection bias.

[10] It is somewhat disconcerting that many $\boldsymbol{z}$ variables in table 7.4 do not have a statistically significant effect on labor market participation and effective wages. To some extent this is caused by multicollinearity of the regional variables. Although multicollinearity does not result in biased estimates, it reduces the precision of the estimated parameters of the correlated regressors, which leads to larger standard errors and thus, to less statistical significance. However, since we are predominantly interested in the effect of the traded quantities of labor on the effective wages, the lack of statistical significance of the $\boldsymbol{z}$ variables has only minor negative consequences on the essential part of our paper.

Table 7.2: Estimation Results of the 2nd-Stage Profit Function

Parameter	$i = c$		$i = a$		$i = v$		$i = L$	
	Coef.	(t-val)	Coef.	(t-val)	Coef.	(t-val)	Coef.	(t-val)
α_i	-31261	(-2.31)	33699	(2.07)	-5480	(-0.37)	-62939	(-6.95)
β_{ic}	53083	(1.86)	64866	(2.75)	-84580	(-2.13)	-33368	(-3.46)
β_{ia}	64866	(2.75)	116773	(2.47)	-168328	(-2.68)	-13311	(-0.63)
β_{iv}	-84580	(-2.13)	-168328	(-2.68)	247344	(2.72)	5564	(0.32)
β_{iL}	-33368	(-3.46)	-13311	(-0.63)	5564	(0.32)	41115	(6.28)
δ_{ig}	6815	(4.59)	303	(0.14)	-6087	(-4.04)	-3181	(-2.81)
δ_{ik}	0.124	(4.40)	0.291	(7.49)	-0.167	(-6.97)	$7.87{\cdot}10^{-3}$	(0.20)
γ_{gg}	-172	(-1.28)						
γ_{gk}	$9.84{\cdot}10^{-3}$	(2.09)						
γ_{kk}	$-3.55{\cdot}10^{-7}$	(-2.26)						
R^2	0.747		0.492		0.821		0.278	

Note: For definitions of the estimated coefficients see equation (7.15), where the subscripts c, a, v, L, g, and k indicate crop products, animal products, variable inputs, labor, land, and capital, respectively. The standard errors of the coefficients are calculated using the bootstrap resampling method (Efron, 1979; Efron and Tibshirani, 1993). Monotonicity is fulfilled at 97.0% of the observations. The R^2 values are almost identical to the model without convexity imposed, indicating that the data do not unreasonably contradict the convexity constraint (see Henning and Henningsen, forthcoming).

The on-farm wage rate increases significantly with hired labor (table 7.4), indicating a convex labor cost function and thus the presence of increasing NTC or heterogeneity. Market imperfections appear more pronounced in on-farm labor markets than in off-farm labor markets. If an average household increases hired labor by 1%, the marginal cost rises by 0.259%. In contrast to the labor supply side, the estimated parameter of the inverse Mill's ratio is significantly different from zero.

We conclude that our estimated FHM is non-separable because the t-tests reject both null hypotheses.

Table 7.3: Estimation Results of the AIDS

Parameter	$i = m$		$i = a$		$i = L$	
	Coef.	(t-val.)	Coef.	(t-val.)	Coef.	(t-val.)
α_i	0.555	(9.86)	0.185	(14.79)	0.260	(4.18)
β_i	-0.170	(-9.15)	-0.031	(-7.36)	0.201	(9.95)
γ_{im}	0.034	(1.28)	0.021	(0.79)	-0.055	(-5.34)
γ_{ia}	0.021	(0.79)	0.010	(0.35)	-0.031	(-9.36)
γ_{iL}	-0.055	(-5.34)	-0.031	(-9.36)	0.086	(7.97)
R^2	0.409		0.585		0.504	

Note: For definitions of the estimated coefficients see equation (7.16), where the subscripts m, a, and L indicate purchased market goods, self-produced goods, and leisure, respectively. The standard errors of the coefficients that have not been directly estimated are calculated with the formula of Klein (1953, p. 258). α_0 is set to 10.8, because this value gives the highest likelihood value of the AIDS Model. Monotonicity is fulfilled at 99.5% of the observations and concavity is fulfilled at 88.4% of the observations.

7.6.2 Elasticities

Given our estimation results, we calculate the full set of price elasticities according to equations (7.13) and (7.14) using sample means. Elasticities for perfect labor markets (separable model) are computed using equation (7.13), setting the second term on the right (the indirect component) equal to zero. Elasticities for imperfect labor markets (non-separable model) are calculated for all four labor market regimes defined in the theoretical section. A detailed derivation of the FHM elasticities is available in Henning and Henningsen (forthcoming).

To assess whether the degree of market imperfection has an impact on farm price responses, we compare the corresponding price elasticities across labor market regimes. The standard errors of the estimated price elasticities and the differences between elasticities derived for different labor market regimes are computed using the formula in Klein (1953, p. 258).

Table 7.5 summarizes the main results and shows the elasticities for three labor market regimes: perfect, imperfect, and missing labor markets. The reader is referred to Henning and Henningsen (forthcoming) for a more comprehensive presentation of the elasticities and their standard errors.

Table 7.4: Estimated Coefficients of Labor Market Equations

	Labor Supply		Labor Demand	
Regressor	1st Step: Probit	2nd Step: 2SLS	1st Step: Probit	2nd Step: 2SLS
Constant	-0.196	103.929 *	3.988	-31.012
X_L^s/T_L		-73.567 **		
X_L^h				0.047 ***
N_k	0.129		0.084	
N_w	0.158 *	-3.496 *	-0.382 ***	
N_o	-0.022	-3.945	-0.296 **	
D_f	0.388	-6.448	-0.199	
A_h	-0.003	2.190 *	-0.119 *	
A_h^2	-7.1·10^{-5}	-0.025 *	1.3·10^{-3} *	
R_g	0.008		0.005	
R_k	-6.5·10^{-7}		1.9·10^{-6} ***	
R_k/R_g	1.0·10^{-5}		-7.2·10^{-6}	2.1·10^{-4} **
P_c	3.091		3.836	
P_a	0.252		-0.115	
P_v	-1.608		-3.906	
N_c	0.142	-1.652	-0.139	4.511
W_u	-0.025	-0.427	-0.010	2.841 **
W_i	-0.030	-0.136	0.001	0.733 *
W_t	-0.007	-0.542	0.012	-0.217
W_r	0.034 **	-0.639 **	-0.030 *	-1.004 **
$\widetilde{P}_L$	-0.014	-0.184	0.005	0.112
IMR Supply		-1.737		
IMR Demand				-15.987 **
ρ	-0.099		-0.099	
R^2		0.307		0.425

Note: *, **, and *** denote statistical significance at the 10%, 5% and 1% level, respectively. Wald test of the joint significance of the exclusion variables: labor supply $\chi^2 = 9.595$, df = 7, p-value = 0.213; labor demand $\chi^2 = 41.531$, df = 11, p-value = 0.00002. Variables: X_L^s = supplied labor [hours], T_L = total time available [hours], X_L^h = hired labor [hours], N_k = number of family members up to 14 years, N_w = number of family members between 15 and 60 years, N_o = number of family members older than 60 years, D_f = sex of the household head (male=0, female=1), A_h = age of the household head, R_g = amount of land of the farm [ha], R_k = amount of capital of the farm [1000 PLZ], P_c = price index of crop products (average=1), P_a = price index of animal products (average=1), P_v = price index of variable inputs (average=1), N_c = number of cars owned by the household, W_u = regional unemployment rate [%], W_i = regional density of the road and railroad network [km/100 km^2], W_t = regional number of telephones per 1,000 population, W_r = proportion of the population that lives in rural areas [%], $\widetilde{P}_L$ = relative average regional wage level (Poland=100), IMR = inverse Mill's ratio.

Table 7.5: Estimated Price Elasticities of Farm Households

	P_c		P_a		P_v		P_m		P_L	
	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)
Perfect labor market (separable model)										
X_c	0.43^a	(1.99)	0.50^a	(2.90)	-0.57^a	(-2.03)	0.00^a		-0.36	(-3.77)
X_a	0.32^a	(2.90)	0.53^a	(2.49)	-0.73^a	(-2.62)	0.00^a		-0.12	(-0.88)
X_v	0.36^a	(2.03)	0.73^a	(2.62)	-1.08^a	(-2.69)	0.00^a		-0.00	(-0.01)
X_L	0.34^a	(3.77)	0.17^a	(0.88)	-0.00^a	(-0.01)	0.00^a		-0.51	(-6.29)
C_m	0.13^a	(6.08)	0.33^a	(3.26)	-0.21^a	(-6.08)	-0.67^a	(-6.80)	0.45	(4.20)
C_a	0.17^a	(7.70)	-0.55^a	(-1.25)	-0.27^a	(-7.70)	0.50^a	(1.20)	0.18	(0.41)
C_L	0.43^a	(42.25)	0.61^a	(39.18)	-0.69^a	(-42.25)	-0.19^a	(-9.46)	-0.07	(-3.22)
X_L^n	-19.15^a	(-13.18)	-22.20^a	(-7.11)	22.00^a	(9.08)	6.16^a	(9.46)	10.30	(7.07)
X_L^f	0.34^a	(3.77)	0.17^a	(0.88)	-0.00^a	(-0.01)	0.00^a		-0.51	(-6.29)
P_L^*	0.00^a		0.00^a		0.00^a		0.00^a		1.00	
Imperfect labor market (non-separable model: supplying and hiring labor)										
X_c	0.28^b	(1.53)	0.33^b	(2.09)	-0.40^b	(-1.55)	0.05^b	(2.56)		
X_a	0.27^a	(2.44)	0.48^a	(2.31)	-0.68^a	(-2.27)	0.02^a	(0.86)		
X_v	0.36^a	(2.10)	0.73^a	(2.57)	-1.08^a	(-2.62)	0.00^a	(0.01)		
X_L	0.14^b	(1.51)	-0.06^b	(-0.41)	0.23^b	(1.84)	0.07^b	(3.11)		
C_m	0.30^b	(5.87)	0.52^b	(4.64)	-0.40^b	(-6.42)	-0.72^b	(-7.28)		
C_a	0.22^b	(7.50)	-0.49^b	(-1.13)	-0.33^b	(-8.20)	0.49^b	(1.17)		
C_L	0.36^b	(15.23)	0.52^b	(17.96)	-0.60^b	(-20.81)	-0.17^b	(-7.99)		
X_L^n	-13.41^a	(-3.12)	-15.55^a	(-3.10)	15.41^a	(3.11)	4.31^a	(3.10)		
X_L^s	-5.47	(-3.46)	-6.34	(-3.28)	6.29	(3.38)	1.76	(3.28)		
X_L^h	1.30	(1.35)	1.51	(1.18)	-1.50	(-1.27)	-0.42	(-1.18)		
X_L^f	0.04^{ab}	(0.14)	-0.19^{ab}	(-0.55)	0.37^{ab}	(1.12)	0.11^{ab}	(1.17)		
P_L^*	0.40^b	(3.60)	0.46^b	(3.39)	-0.46^b	(-3.51)	-0.13^b	(-3.40)		
Missing labor market (non-separable model: autarkic in labor)										
X_c	-0.07^c	(-0.51)	-0.07^c	(-0.49)	0.00^c	(0.01)	0.16^c	(3.40)		
X_a	0.16^a	(0.88)	0.35^a	(1.32)	-0.55^a	(-1.44)	0.05^a	(0.92)		
X_v	0.35^a	(1.84)	0.72^a	(2.24)	-1.08^a	(-2.37)	0.00^a	(0.01)		
X_L	-0.35^c	(-5.77)	-0.63^c	(-9.15)	0.80^c	(11.88)	0.22^c	(6.00)		
C_m	0.69^c	(10.50)	0.97^c	(8.51)	-0.85^c	(-11.88)	-0.85^c	(-8.75)		
C_a	0.35^c	(5.16)	-0.34^c	(-0.82)	-0.48^c	(-6.07)	0.44^c	(1.08)		
C_L	0.17^c	(5.77)	0.31^c	(9.15)	-0.39^c	(-11.88)	-0.11^c	(-6.00)		
X_L^n	0.00^b		0.00^b		0.00^b		0.00^b			
X_L^f	-0.35^b	(-5.77)	-0.63^b	(-9.15)	0.80^b	(11.88)	0.22^b	(6.00)		
P_L^*	1.36^c	(9.17)	1.58^c	(9.44)	-1.56^c	(-9.79)	-0.44^c	(-5.65)		

Note: Variables: $X_.$ = netput quantities, $C_.$ = consumed quantities, $P_.$ = exogenous prices, $P_.^*$ = endogenous shadow prices; subscripts: c = crop products, a = animal products, v = variable inputs, L = labor/leisure; superscripts of X_L (labor quantities): f = family labor on the farm, h = hired, s = supplied, n = net supplied. For each specific elasticity the values that have a common alphabetic character do not differ significantly. For instance, the elasticity of X_c with respect to P_c has different letters for all three types of labor market imperfections, which means that these three values significantly differ. On the other hand, the elasticity of X_c with respect to P_a has the same letter for all three types of labor market imperfections, which means that these three values do not differ significantly.

Overall, we observe mixed results. For all consumer goods, crop products, and farm labor input, the degree of labor market imperfection has a significant influence on price responses. By contrast, price elasticities for animal products and variable inputs do not significantly differ across labor market regimes, indicating that the degree of market imperfection has only a negligible impact on household's price responses.

How can these results be explained? According to equation (7.13), for any good Q and any exogenous price P_j, $j \in \{c, a, v, m\}$, the difference in price elasticities between perfect and imperfect/missing labor market regimes equals $(\partial Q/\partial P_L^*)\,(P_L^*/Q) \cdot (\mathrm{d}\,P_L^*/\mathrm{d}\,P_j)\,(P_j/P_L^*)$. The first term denotes the cross-price elasticity for good Q with respect to the wage rate and the second term is the shadow price elasticity. The latter measures the impact of an exogenous price change on the shadow price of labor, while the first measures the change of the consumed or produced quantity of good Q induced by a change in the shadow price of labor.

Differences in price elasticities can thus result from either high cross-price elasticities or high shadow price elasticities, or both. Relatively high cross-price elasticities are observed for crop products (-0.36), farm labor input (-0.51) and purchased consumer goods (0.45) (see table 7.5). For these goods, we also observe the largest and statistically significant differences in price elasticities across market regimes. High shadow price elasticities were obtained for missing markets, while low values were found for imperfect labor markets. This reinforces our finding that the degree of imperfection due to NTC or heterogeneity is moderate. Among all commodity prices, the one for purchased consumer goods (P_m) has the lowest impact on the shadow price for labor, as can be seen from the right-hand column of table 7.5. This can be explained with reference to equation (14), where the numerator captures the commodity specific income and substitution effects. The lower these effects, the lower are the shadow price elasticities.

Table 7.5 also shows that adjustments of net labor supply ($X_L^n = X_L^s - X_L^h$) do not differ significantly between perfect and imperfect labor markets. However, for both regimes, these adjustments differ significantly from zero. Of course, labor adjustment is zero for missing markets.

Finally, in the Polish case, market imperfection reduces household's responses to exogenous price changes on the production side, i.e. most price elasticities decrease in absolute terms with the degree of market imperfection. For example, for perfect labor markets crop output and farm labor input show a clear positive response with respect to increased crop prices. These responses are significantly smaller if labor markets are imperfect, and become negative in missing labor markets, implying even an inverse supply response.

7.7 Conclusion

This article developed a farm household model that incorporates labor market imperfections due to fixed (FTC), proportional (PTC), and non-proportional variable (NTC) transaction costs as well as heterogeneity in on-farm and off-farm labor markets. In contrast to existing studies that incorporate only FTC and PTC, the model developed here allows for non-separability, even when households buy or sell labor.

Comparative static analysis indicates that price responses deviate from perfect labor markets, even when the household buys or sells labor, if NTC or labor heterogeneity exist. Furthermore, price elasticities in imperfect labor markets generally lie between the corresponding elasticities in absent and perfect labor markets.

The model also provides a quantitative measure of the degree of market imperfection due to NTC and heterogeneity, and allows for a test of whether NTC and heterogeneity can be excluded from the estimation without loss of explanatory power.

Applying the model to farm household data from Mid-West Poland shows that NTC and heterogeneity play a significant role in explaining households' behavior. However, in the Polish case, market imperfection due to NTC or heterogeneity is rather moderate, with the effect of NTC and heterogeneity more pronounced when hiring on-farm labor than supplying off-farm labor. Econometric estimation of our generalized FHM approach is rather cumbersome, because we have to control simultaneously for various possible endogeneity and selectivity biases. Therefore, the question arises if this more complex model is worth the effort. From the perspective of policy makers, we must ask whether incorporating NTC and heterogeneity provides estimates of elasticities that are quite different from what could have been obtained otherwise. Here our analysis delivers mixed results. While differences are statistically significant and are considerable for all consumer and most producer goods, they are not for animal products and variable inputs.

Acknowledgment

The authors thank Wade Brorsen and three anonymous reviewers for their very helpful comments and suggestions. We also thank Awudu Abdulai, Uwe Latacz-Lohmann, and Jerzy Michalek for very helpful comments. Any remaining errors are the authors'.

Supplementary Appendix

Note: This supplementary appendix will be published in the "AgEcon Search" online library (http://agecon.lib.umn.edu). In the main article this supplementary appendix is cited as "Henning and Henningsen (forthcoming)".

7.A1 Motivation of the Labor Market Model

7.A1.1 Simultaneous Demand of On-Farm Labor and Supply of Off-Farm Labor

Simultaneously demanding on-farm labor and supplying off-farm labor can be rational with a strictly convex labor cost function and a strictly concave labor income function. To observe this, assume that in autarky the shadow price of labor on the farm would be lower than the marginal revenue of selling off-farm labor and higher than the marginal cost of hiring on-farm labor. Obviously, under this assumption, utility maximizing implies that the farm household supplies off-farm labor until marginal revenue equals the shadow price of labor, while the household demands on-farm labor until marginal cost equals the shadow price of labor or hired labor equals optimal labor input, i.e. the household no longer works on its own farm. Now, given strict convex and strict concave labor costs and income functions, there always exists an interior solution, i.e. the household simultaneously supplies and demands labor and works on its own farm. For instance, if the skills of the household members to work off-farm are very heterogeneous, it is rational to simultaneously supply high-priced labor of well-educated household members and hire cheap agricultural labor (see also Sadoulet, de Janvry, and Benjamin, 1996).

7.A1.2 Examples of Non-proportional Variable Transaction Costs

In this section we provide some intuitive examples of non-proportional variable transaction costs (NTC). It is well recognized in the literature that participation in rural labor markets is often plagued by adverse selection and moral hazard problems due to asymmetric information regarding the quality of the labor force (Eswaran and Kotwal, 1986; Spence, 1976) and the effort of hired labor, respectively (Frisvold, 1994; Sadoulet, de Janvry, and Benjamin, 1998; Eswaran and Kotwal, 1986). Generally, moral hazard and adverse selection problems might change non-proportionally with the quantity of traded goods, implying NTC for both on-farm labor demand and off-farm labor supply. Theoretically, it is unclear how these costs vary, i.e. if they are increasing, decreasing, or proportional to the amount of hired or supplied labor.

For example, in the case of moral hazard problems of hired on-farm labor, it is well recognized that employers cannot easily infer labor effort indirectly by observing final output, due to the stochastic and seasonal nature of agricultural production. Therefore, supervision costs rise to control for moral hazard problems (Frisvold, 1994; Feder, 1985). Marginal costs to supervise hired labor may increase along with the units of hired labor due to an increase in the probability of free-riding, the greater importance of coordinating work inputs, and the increased effort to control for social conflicts among employees.

Moreover, adverse selection problems due to asymmetric information on the quality of hired labor might lead to transaction costs in rural labor markets. These transaction costs might be partially reduced by adequate formal institutions (Spence, 1976). However, in rural labor markets, adequate formal institutions that avoid adverse selection problems, e.g. formal education certificates, are often incompletely developed. In that case, a firm might use informal screening mechanisms to learn about the quality of workers, e.g. information from peer groups or rural organizations (Granovetter, 1973; Sadoulet, de Janvry, and Benjamin, 1998). Accessability to peer groups or rural organizations varies, i.e. workers living in the neighborhood might have more access than those living in a more distant village. Thus, the potential to control for adverse selection problems increases when firms shift their demand from local to regional labor markets, implying increasing marginal NTC.

Moreover, even if information, search, and bargaining costs are considered as fixed costs, they occur for each labor contract. Therefore, from the perspective of the farm household, total costs, including all labor contracts, are no longer fixed costs but vary with the number of workers. Finally, other transaction costs might also vary with the number of labor contracts, e.g. there are only slight additional costs if one or two people travel to the city in the same car or family members who work for the same firm might reduce search and bargaining costs for succeeding family workers. On the other hand, some part-time jobs might be available near the farm, while full-time jobs are only available in larger settlements farther away, implying increasing transportation costs.

7.A2 Theoretical Results

Table 7.A1: Theoretical Effects of Exogenous Price Changes

Behavior	Variable	Non-separable Model				Separable Model				
		P_c	P_a	P_v	P_m	P_c	P_a	P_v	P_L	P_m
Farm	X_c	?	?	?	?	+	?	(-)	(-)	0
	X_a	?	?	?	?	?	+	(-)	(-)	0
	$\|X_v\|$	?	?	?	?	(+)	(+)	-	(-)	0
	$\|X_L\|$	?	?	?	?	(+)	(+)	(-)	-	0
Consumption	C_m	(+)	(+)	(-)	?	(+)	(+)	(-)	(+)	(-)
	C_a	(+)	?	(-)	?	(+)	?	(-)	(+)	?
	C_L	?	?	?	?	(+)	(+)	(-)	?	?
Labor market	X_L^n	(-)	(-)	(+)	?	(-)	(-)	(+)	(+)	?
	X_L^s	(-)	(-)	(+)	?					
	X_L^h	(+)	(+)	(-)	?					
	P_L^*	(+)	(+)	(-)	?					

Note: It is assumed that goods are not inferior, technologies are not regressive, and households are net suppliers of labor and self-produced agricultural goods.

Variables: $X_.$ = netput quantities, $C_.$ = consumed quantities, $P_.$ = exogenous prices, $P_.^*$ = endogenous shadow prices; subscripts: c = crop products, a = animal products, v = variable inputs, L = labor/leisure; superscripts of X_L (labor quantities): h = hired, s = supplied, n = net supplied.

Symbols indication the direction of the effects:

0	=	clear, no effect;
+/-	=	clear, increase/decrease;
(+)/(-)	=	unclear, but most likely an increase/decrease (assuming labor and variable inputs are complements, and consumption goods are net-substitutes);
?	=	unclear.

7.A3 Symmetric Normalized Quadratic (SNQ) Profit Function

This functional form is also traded under the name of "symmetric generalized McFadden function" (Diewert and Wales, 1992).

7.A3.1 First Stage Profit Function

We follow Lopez (1984) and determine the shadow price of labor on the farm by estimating a profit function assuming constant returns to labor. In this case a symmetric normalized quadratic (SNQ) profit function (Diewert and Wales, 1987, 1992; Kohli, 1993) has following form:

$$\Pi\left(\boldsymbol{p}_{pn}, \boldsymbol{r}_n, X_{Ln}\right) = X_{Ln} \left(\begin{array}{l} \sum\limits_{i \in \{c,a,v\}} \alpha_i P_{in} + \frac{1}{2} w_n^{-1} \sum\limits_{i \in \{c,a,v\}} \sum\limits_{j \in \{c,a,v\}} \beta_{ij} P_{in} P_{jn} \\ + \sum\limits_{i \in \{c,a,v\}} \sum\limits_{j \in \{g,k\}} \delta_{ij} P_{in} R_{jn} + \frac{1}{2} w_n \sum\limits_{i \in \{g,k\}} \sum\limits_{j \in \{g,k\}} \gamma_{ij} R_{in} R_{jn} \end{array} \right) \tag{7.34}$$

where n indicates the observation (household), Π is the profit function, X_{Ln} is the labor deployed on the farm, $w_n = \sum_{i \in \{c,a,v\}} \theta_i P_{in}$ is a factor to normalize prices, $\theta_i = \sum_n P_{in} |X_{in}| / \sum_n \sum_{j \in \{c,a,v\}} P_{jn} |X_{jn}|$; $i \in \{c, a, v\}$ are predetermined weights of the individual netput prices, $\boldsymbol{p}_{pn} = (P_{an}, P_{cn}, P_{vn})$ indicates the netput prices, X_{in}; $i \in \{c, a, v\}$ denotes the quantity indices of the netputs, $\boldsymbol{r}_n = (R_{gn}, R_{kn})$ represents the quasi-fixed factors land (R_g) and capital (R_k), and α_i, β_{ij}, δ_{ij}, and γ_{ij} are the parameters to be estimated. To identify all β_{ij}, we impose the restrictions $\sum_{j \in \{c,a,v\}} \beta_{ij} \overline{P_j} = 0$; $i \in \{c, a, v\}$, where $\overline{P_j}$ are the mean prices (Diewert and Wales, 1987, p. 54).

The corresponding netput equations can be obtained using Hotelling's Lemma:

$$X_{in}\left(\boldsymbol{p}_{pn}, \boldsymbol{r}_n, X_{Ln}\right) = \frac{\partial \Pi\left(\boldsymbol{p}_{pn}, \boldsymbol{r}_n, X_{Ln}\right)}{\partial P_{in}} \tag{7.35}$$

$$= X_{Ln} \left(\begin{array}{l} \alpha_i + w_n^{-1} \sum\limits_{j \in \{c,a,v\}} \beta_{ij} P_{jn} \\ -\frac{1}{2} \theta_i w_n^{-2} \sum\limits_{j \in \{c,a,v\}} \sum\limits_{k \in \{c,a,v\}} \beta_{jk} P_{jn} P_{kn} \\ + \sum\limits_{j \in \{g,k\}} \delta_{ij} R_{jn} + \frac{1}{2} \theta_i \sum\limits_{j \in \{g,k\}} \sum\limits_{k \in \{g,k\}} \gamma_{jk} R_{jn} R_{kn} \end{array} \right) \tag{7.36}$$

7.A3.2 Second Stage Profit Function

At the second stage we estimate a symmetric normalized quadratic (SNQ) profit function (Diewert and Wales, 1987, 1992; Kohli, 1993) with labor as variable input:

$$\begin{aligned}\Pi\left(\boldsymbol{p}_{pn},\boldsymbol{r}_n\right) &= \sum_{i\in\{c,a,v,L\}}\alpha_i P_{in}+\frac{1}{2}w_n^{-1}\sum_{i\in\{c,a,v,L\}}\sum_{j\in\{c,a,v,L\}}\beta_{ij}P_{in}P_{jn} \\ &+\sum_{i\in\{c,a,v,L\}}\sum_{j\in\{g,k\}}\delta_{ij}P_{in}R_{jn}+\frac{1}{2}w_n\sum_{i\in\{g,k\}}\sum_{j\in\{g,k\}}\gamma_{ij}R_{in}R_{jn}\end{aligned} \tag{7.37}$$

where $w_n = \sum_{i\in\{c,a,v,L\}}\theta_i P_{in}$ is a factor to normalize prices, $\theta_i = \sum_n P_{in}\left|X_{in}\right|/\sum_n\sum_{j\in\{c,a,v,L\}} P_{jn}\left|X_{jn}\right|$; $i\in\{c,a,v,L\}$ are predetermined weights of the individual netput prices, $\boldsymbol{p}_{pn}=(P_{an},P_{cn},P_{vn},P_{Ln})$ indicates the netput prices, X_{in}; $i\in\{c,a,v,L\}$ denotes the quantity indices of the netputs, $\boldsymbol{r}_n=(R_{gn},R_{kn})$ represents the quasi-fixed factors land (R_g) and capital (R_k), and α_i, β_{ij}, δ_{ij}, and γ_{ij} are the parameters to be estimated. To identify all β_{ij}, we impose the restrictions $\sum_{j\in\{c,a,v,L\}}\beta_{ij}\overline{P_j}=0$; $i\in\{c,a,v,L\}$, where $\overline{P_j}$ are the mean prices (Diewert and Wales, 1987, p. 54).

The corresponding netput equations can be obtained using Hotelling's Lemma:

$$X_{in}\left(\boldsymbol{p}_{pn},\boldsymbol{r}_n\right) = \frac{\partial\Pi\left(\boldsymbol{p}_{pn},\boldsymbol{r}_n\right)}{\partial P_{in}} \tag{7.38}$$

$$\begin{aligned}&= \alpha_i + w_n^{-1}\sum_{j\in\{c,a,v,L\}}\beta_{ij}P_{jn}-\frac{1}{2}\theta_i w_n^{-2}\sum_{j\in\{c,a,v,L\}}\sum_{k\in\{c,a,v,L\}}\beta_{jk}P_{jn}P_{kn} \\ &+\sum_{j\in\{g,k\}}\delta_{ij}R_{jn}+\frac{1}{2}\theta_i\sum_{j\in\{g,k\}}\sum_{k\in\{g,k\}}\gamma_{jk}R_{jn}R_{kn}\end{aligned} \tag{7.39}$$

7.A4 Labor Market Analysis

7.A4.1 Labor Supply

To estimate the marginal revenue of supplying labor, we assume the following specifications of the average regional wage level $\overline{P}_L$, the household-specific wage shifters b^s, and the variable transaction costs TC_v^s, which include proportional (PTC) and non-proportional variable (NTC) transaction costs:

$$\overline{P}_L = \widetilde{P}_L\beta_p^s \tag{7.40}$$

$$b^s\left(X_L^s,\boldsymbol{z}_L^s\right) = \beta_0^s+\boldsymbol{z}_L^{s\prime}\boldsymbol{\beta}_L^s+X_L^s\beta_{L1}^s \tag{7.41}$$

$$TC_v^s\left(X_L^s,\boldsymbol{z}_v^s\right) = \left(\boldsymbol{z}_v^{s\prime}\boldsymbol{\beta}_v^s\right)X_L^s+\beta_{v1}^s{X_L^s}^2 \tag{7.42}$$

where $\widetilde{P}_L$ is a proxy for the average regional wage level. The specification of b^s shows that $\boldsymbol{z}_L^{s\prime}\boldsymbol{\beta}_L^s$ indicates general wage differences between the households, while $X_L^s\beta_{L1}^s$ refers to a wage shift due to a changing amount of supplied labor, which is caused by heterogeneity within each household. The specification of TC_v^s is derived from a second-order Taylor series approximation of the true transaction costs (see section 7.A4.3). It shows that $\boldsymbol{z}_v^{s\prime}\boldsymbol{\beta}_v^s$ denotes proportional transaction costs per unit of labor, and $\beta_{v1}^s X_L^{s\,2}$ are non-proportional variable transaction costs.

Substituting these specifications into equation (7.5) of the main article, we get the empirical specification used for the estimation, which is presented in equation (7.18) of the main article.

$$
\begin{aligned}
P_L^s &= \overline{P}_L + b^s\left(X_L^s, \boldsymbol{z}_L^s\right) - \frac{\partial TC_v^s\left(X_L^s, \boldsymbol{z}_v^s\right)}{\partial X_L^s} && (7.43)\\
&= \widetilde{P}_L\beta_p^s + \beta_0^s + \boldsymbol{z}_L^{s\prime}\boldsymbol{\beta}_L^s + X_L^s\beta_{L1}^s - \boldsymbol{z}_v^{s\prime}\boldsymbol{\beta}_v^s - 2X_L^s\beta_{v1}^s && (7.44)\\
&= \beta_0^s + \widetilde{P}_L\beta_p^s + \boldsymbol{z}_L^{s\prime}\boldsymbol{\beta}_L^s - \boldsymbol{z}_v^{s\prime}\boldsymbol{\beta}_v^s + X_L^s(\beta_{L1}^s - 2\beta_{v1}^s) && (7.45)\\
&= \beta_0^s + \boldsymbol{z}^{s\prime}\boldsymbol{\beta}^s + X_L^s\beta_1^s && (7.46)
\end{aligned}
$$

with $\boldsymbol{z}^s = \left(\widetilde{P}_L, \boldsymbol{z}_L^{s\prime}, \boldsymbol{z}_v^{s\prime}\right)'$, $\boldsymbol{\beta}^s = \left(\beta_p^s, \boldsymbol{\beta}_L^{s\prime}, -\boldsymbol{\beta}_v^{s\prime}\right)'$, and $\beta_1^s = \beta_{L1}^s - 2\beta_{v1}^s$.

Neglecting FTC, we can derive the net off-farm labor revenue function f from the estimated coefficients of equation (7.46) by applying equation (7.6) of the main article:

$$
f\left(X_L^s\right) + TC_f^s = \int_0^{X_L^s}\left(\widehat{\beta}_0^s + \boldsymbol{z}^{s\prime}\widehat{\boldsymbol{\beta}}^s + X_L^s\widehat{\beta}_1^s\right)\mathrm{d}\,X_L^s = \left(\widehat{\beta}_0^s + \boldsymbol{z}^{s\prime}\widehat{\boldsymbol{\beta}}^s\right)X_L^s + \frac{1}{2}\widehat{\beta}_1^s\,X_L^{s\,2} \qquad (7.47)
$$

7.A4.2 Labor Demand

To estimate the marginal cost of hiring labor, we assume the following specifications of the average regional wage level $\overline{P}_L$, the farm-specific wage shifters b^h, and the variable transaction costs TC_v^h, which include PTC and NTC:

$$
\begin{aligned}
\overline{P}_L &= \widetilde{P}_L\beta_p^h && (7.48)\\
b^h\left(X_L^h, \boldsymbol{z}_L^h\right) &= \beta_0^h + \boldsymbol{z}_L^{h\prime}\boldsymbol{\beta}_L^h + X_L^h\beta_{L1}^h && (7.49)\\
TC_v^h\left(X_L^h, \boldsymbol{z}_v^h\right) &= \left(\boldsymbol{z}_v^{h\prime}\boldsymbol{\beta}_v^h\right)X_L^h + \beta_{v1}^h X_L^{h2} && (7.50)
\end{aligned}
$$

where $\widetilde{P}_L$ is a proxy for the average regional wage level. The specification of b^h shows that $\boldsymbol{z}_L^{h\prime}\boldsymbol{\beta}_L^h$ indicates general wage differences between the farms, while $X_L^h\beta_{L1}^h$ refers to a wage shift due to a changing amount of hired labor, which is caused by heterogeneity within the hired workers of each farm. The specification of TC_v^h is derived from a second-order Taylor series approximation of the true transaction costs (see section 7.A4.3). It shows that $\boldsymbol{z}_v^{h\prime}\boldsymbol{\beta}_v^h$ denotes proportional transaction costs per unit of labor, and $\beta_{v1}^h X_L^{h\,2}$ are non-proportional transaction costs.

Substituting these specifications into equation (7.7) of the main article, we get the empirical specification used for the estimation, which is presented in equation (7.19) of the main article.

$$
\begin{aligned}
P_L^h &= \overline{P}_L + b^h\left(X_L^h, \boldsymbol{z}_L^h\right) + \frac{\partial TC_v^h\left(X_L^h, \boldsymbol{z}_v^h\right)}{\partial X_L^h} && (7.51)\\
&= \widetilde{P}_L\beta_p^h + \beta_0^h + \boldsymbol{z}_L^{h\prime}\boldsymbol{\beta}_L^h + X_L^h\beta_{L1}^h + \boldsymbol{z}_v^{h\prime}\boldsymbol{\beta}_v^h + 2X_L^h\beta_{v1}^h && (7.52)\\
&= \beta_0^h + \widetilde{P}_L\beta_p^h + \boldsymbol{z}_L^{h\prime}\boldsymbol{\beta}_L^h + \boldsymbol{z}_v^{h\prime}\boldsymbol{\beta}_v^h + X_L^h(\beta_{L1}^h + 2\beta_{v1}^h) && (7.53)\\
&= \beta_0^h + \boldsymbol{z}^{h\prime}\boldsymbol{\beta}^h + X_L^h\beta_1^h && (7.54)
\end{aligned}
$$

with $\boldsymbol{z}^h = \left(\widetilde{P}_L, \boldsymbol{z}_L^{h\prime}, \boldsymbol{z}_v^{h\prime}\right)'$, $\boldsymbol{\beta}^h = \left(\beta_p^h, \boldsymbol{\beta}_L^{h\prime}, \boldsymbol{\beta}_v^{h\prime}\right)'$, and $\beta_1^h = \beta_{L1}^h + 2\beta_{v1}^h$.

Neglecting FTC, we can derive the effective cost function for hired labor g from the estimated coefficients of equation (7.54) by applying equation (7.8) of the main article:

$$
g\left(X_L^h\right) - TC_f^h = \int_0^{X_L^h}\left(\widehat{\beta}_0^h + \boldsymbol{z}^{h\prime}\widehat{\boldsymbol{\beta}}^h + \widehat{\beta}_1^h X_L^h\right) \mathrm{d}\, X_L^h = \left(\widehat{\beta}_0^h + \boldsymbol{z}^{h\prime}\widehat{\boldsymbol{\beta}}^h\right) X_L^h + \frac{1}{2}\widehat{\beta}_1^h\, X_L^{h\,2} \quad (7.55)
$$

7.A4.3 Second-order Taylor Series Approximation of Variable Transaction Costs

We assume that the transactions costs (TC) are a function of the traded quantity (X_L) and some further factors that influence variable transaction costs ($\boldsymbol{z}_v$)[11]:

$$
TC = f\begin{pmatrix} X_L \\ \boldsymbol{z}_v \end{pmatrix} \quad (7.56)
$$

where X_L is a scalar that represents X_L^s or X_L^h and $\boldsymbol{z}_v$ is a vector that represents $\boldsymbol{z}_v^s$ or $\boldsymbol{z}_v^h$.

[11] In this section we ignore factors that influence fixed transaction costs because we are interested only in variable transaction costs here.

We approximate the true transaction costs at point $\begin{pmatrix} X_L^0 \\ \boldsymbol{z}_v^0 \end{pmatrix}$ by a second-order Taylor series[12]:

$$
\begin{aligned}
TC^* &= f\begin{pmatrix} X_L^0 \\ \boldsymbol{z}_v^0 \end{pmatrix} + \begin{pmatrix} X_L - X_L^0 \\ \boldsymbol{z}_v - \boldsymbol{z}_v^0 \end{pmatrix}' \begin{pmatrix} \dfrac{\partial TC}{\partial X_L} \\ \dfrac{\partial TC}{\partial \boldsymbol{z}_v} \end{pmatrix} && (7.57)\\
&\quad + \begin{pmatrix} X_L - X_L^0 \\ \boldsymbol{z}_v - \boldsymbol{z}_v^0 \end{pmatrix}' \begin{pmatrix} \dfrac{\partial^2 TC}{\partial X_L^2} & \dfrac{\partial^2 TC}{\partial X_L \partial \boldsymbol{z}_v} \\ \dfrac{\partial^2 TC}{\partial X_L \partial \boldsymbol{z}_v} & \dfrac{\partial^2 TC}{\partial \boldsymbol{z}_v^2} \end{pmatrix} \begin{pmatrix} X_L - X_L^0 \\ \boldsymbol{z}_v - \boldsymbol{z}_v^0 \end{pmatrix} \\
&= f\begin{pmatrix} X_L^0 \\ \boldsymbol{z}_v^0 \end{pmatrix} + \left(X_L - X_L^0\right) \frac{\partial TC}{\partial X_L} + \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right)' \frac{\partial TC}{\partial \boldsymbol{z}_v} + \left(X_L - X_L^0\right)^2 \frac{\partial^2 TC}{\partial X_L^2} && (7.58)\\
&\quad +2\left(X_L - X_L^0\right) \frac{\partial^2 TC}{\partial X_L \partial \boldsymbol{z}_v} \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right) + \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right)' \frac{\partial^2 TC}{\partial \boldsymbol{z}_v^2} \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right) \\
&= f\begin{pmatrix} X_L^0 \\ \boldsymbol{z}_v^0 \end{pmatrix} + \frac{\partial TC}{\partial X_L} X_L - \frac{\partial TC}{\partial X_L} X_L^0 + \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right)' \frac{\partial TC}{\partial \boldsymbol{z}_v} && (7.59)\\
&\quad + \frac{\partial^2 TC}{\partial X_L^2} X_L^2 - 2\frac{\partial^2 TC}{\partial X_L^2} X_L X_L^0 + \frac{\partial^2 TC}{\partial X_L^2} {X_L^0}^2 + 2\frac{\partial^2 TC}{\partial X_L \partial \boldsymbol{z}_v} \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right) X_L \\
&\quad -2\frac{\partial^2 TC}{\partial X_L \partial \boldsymbol{z}_v} \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right) X_L^0 + \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right)' \frac{\partial^2 TC}{\partial \boldsymbol{z}_v^2} \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right)
\end{aligned}
$$

All terms that do not vary with X_L are considered as fixed transaction costs:

$$
\begin{aligned}
TC_f^* &= f\begin{pmatrix} X_L^0 \\ \boldsymbol{z}_v^0 \end{pmatrix} - \frac{\partial TC}{\partial X_L} X_L^0 + \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right)' \frac{\partial TC}{\partial \boldsymbol{z}_v} + \frac{\partial^2 TC}{\partial X_L^2} {X_L^0}^2 && (7.60)\\
&\quad -2\frac{\partial^2 TC}{\partial X_L \partial \boldsymbol{z}_v} \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right) X_L^0 + \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right)' \frac{\partial^2 TC}{\partial \boldsymbol{z}_v^2} \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0\right)
\end{aligned}
$$

[12] All derivatives are evaluated at point $\begin{pmatrix} X_L^0 \\ \boldsymbol{z}_v^0 \end{pmatrix}$ but in the following this is omitted for better readability.

Now we get for the variable transaction costs

$$
\begin{aligned}
TC_v^* &= TC^* - TC_f^* && (7.61)\\
&= \frac{\partial TC}{\partial X_L} X_L + \frac{\partial^2 TC}{\partial X_L^2} X_L^2 - 2 \frac{\partial^2 TC}{\partial X_L^2} X_L X_L^0 + 2 \frac{\partial^2 TC}{\partial X_L \partial \boldsymbol{z}_v} \left(\boldsymbol{z}_v - \boldsymbol{z}_v^0 \right) X_L && (7.62)\\
&= \left(\frac{\partial TC}{\partial X_L} - 2 \frac{\partial^2 TC}{\partial X_L^2} X_L^0 - 2 \frac{\partial^2 TC}{\partial X_L \partial \boldsymbol{z}_v} \boldsymbol{z}_v^0 + 2 \frac{\partial^2 TC}{\partial X_L \partial \boldsymbol{z}_v} \boldsymbol{z}_v \right) X_L + \frac{\partial^2 TC}{\partial X_L^2} X_L^2 && (7.63)\\
&= \left(\widetilde{\boldsymbol{z}}_v' \boldsymbol{\beta}_v \right) X_L + \beta_{v1} X_L^2 && (7.64)
\end{aligned}
$$

with

$$
\begin{aligned}
\widetilde{\boldsymbol{z}}_v &= \begin{pmatrix} 1 \\ \boldsymbol{z}_v \end{pmatrix} && (7.65)\\
\boldsymbol{\beta}_v &= \begin{pmatrix} \frac{\partial TC}{\partial X_L} - 2 \frac{\partial^2 TC}{\partial X_L^2} X_L^0 - 2 \frac{\partial^2 TC}{\partial X_L \partial \boldsymbol{z}_v} \boldsymbol{z}_v^0 \\ 2 \frac{\partial^2 TC}{\partial X_L \partial \boldsymbol{z}_v} \end{pmatrix} && (7.66)\\
\beta_{v1} &= \frac{\partial^2 TC}{\partial X_L^2} && (7.67)
\end{aligned}
$$

7.A4.4 Exclusion Variables

In a two-step Heckman estimation, the variables that are regressors in the first-step selection equation (say, $\boldsymbol{x}_1$) but are not regressors in the second-step regression equation (say, $\boldsymbol{x}_2$) are called "exclusion variables." If there are no exclusion variables ($\boldsymbol{x}_1 \subseteq \boldsymbol{x}_2$), the sample correction term in the second step (say, λ) is likely to be highly correlated with the other regressors in $\boldsymbol{x}_2$ because λ is a (non-linear) function of a linear combination of the variables in $\boldsymbol{x}_1$ ($\lambda = \phi(\boldsymbol{x}_1{}'\boldsymbol{\gamma})/\Phi(\boldsymbol{x}_1{}'\boldsymbol{\gamma})$, where $\boldsymbol{\gamma}$ are the coefficients of the selection equation and ϕ and Φ are probability density function (pdf) and the cumulative distribution function (cdf) of the standard normal distribution, respectively). Hence, the purpose of exclusion variables is to reduce the correlation among the regressors (multicollinearity) in the second-step estimation. Although high multicollinearity does not result in biased estimates, it leads to large standard errors, which means that the estimates are rather imprecise.

The exclusion variables for the equations explaining the shadow price of labor can be identified from table 7.4 in the main article. The exclusion variables for the marginal revenue of labor supply (equation (7.24) in the main article) are the number of kids (N_k), land and capital endowment of the farm (R_g, R_k); the capital intensity on the farm (R_k/R_g); and the prices of farm netputs (P_c, P_a, P_v). The exclusion variables for the marginal cost of labor demand (equation (7.25) in the main article) are the age pattern of the household (N_k, N_w, N_o); sex, age, and age squared of the head of the household (D_f, A_h, A_h^2); land and capital endowment of the farm (R_g, R_k); and the prices of farm netputs (P_c, P_a, P_v).

The exclusion variables for the equations explaining the quantity of supplied labor (equations (7.26) and (7.27) in the main article) are variables that are in $\boldsymbol{z}$ but not in $\boldsymbol{z}_x^b$ and $\boldsymbol{z}_x^s$, respectively. The exclusion variables for the equations explaining the quantity of hired labor (equations (7.28) and (7.29) in the main article) are variables that are in $\boldsymbol{z}$ but not in $\boldsymbol{z}_x^b$ and $\boldsymbol{z}_x^h$, respectively. Theoretically, the exclusion variables in (7.26) and (7.28) are the variables that are in $\boldsymbol{z}_f^s$ or $\boldsymbol{z}_f^h$ but not in $\boldsymbol{z}^\pi$, $\boldsymbol{z}^u$, $\boldsymbol{z}^s$, or $\boldsymbol{z}^h$, the exclusion variables in (7.27) are the variables that are in $\boldsymbol{z}_f^s$, $\boldsymbol{z}_f^h$ or $\boldsymbol{z}^h$ but not in $\boldsymbol{z}^\pi$, $\boldsymbol{z}^u$ or $\boldsymbol{z}^s$, and the exclusion variables in (7.29) are the variables that are in $\boldsymbol{z}_f^s$, $\boldsymbol{z}_f^h$ or $\boldsymbol{z}^s$ but not in $\boldsymbol{z}^\pi$, $\boldsymbol{z}^u$ or $\boldsymbol{z}^h$. However, in practice, our data set does not include any variables that influence fixed transaction costs ($\boldsymbol{z}_f^s$, $\boldsymbol{z}_f^h$) but do not influence variable transaction costs or the average skill level ($\boldsymbol{z}^s$, $\boldsymbol{z}^h$). Thus, given the specification of the $\boldsymbol{z}$ variables in section “Data and Empirical Results” in the main article, we have an exclusion variable only in (7.27) (R_K/R_g) but not in the other three X equations. Although this leads to multicollinearity, it does not matter in our special case because we are interested in the fitted values but not the estimated coefficients. As long as multicollinearity is not so high that it rules out estimation, we can calculate fitted values that are orthogonal to the error terms of the estimations of the shadow price of labor (given that the regressors are not correlated with these error terms, too).

7.A5 Assumptions about Error Terms

We assume that the residuals of the participation equations (7.22, 7.23) in the main article, ε^s and ε^h, have a bivariate normal distribution:

$$\begin{pmatrix} \varepsilon^s \\ \varepsilon^h \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & \rho \\ & 1 \end{bmatrix}\right) \tag{7.68}$$

Further, we assume a joint normal distribution of ε^s, ε^h, $\widetilde{\nu}^s$ and $\widetilde{\nu}^h$ with covariances $\sigma^s = cov(\widetilde{\nu}^s, \varepsilon^s)$ and $\sigma^h = cov(\widetilde{\nu}^h, \varepsilon^h)$, where $\widetilde{\nu}^s$ and $\widetilde{\nu}^h$ would be the error terms of equations (7.24) and (7.25) in the main article, respectively, without selectivity terms. From this we can obtain the conditional expectation of the error terms

$$E\left[\widetilde{\nu}^s | Y^{s*} > 0\right] = \sigma^s \lambda^s \tag{7.69}$$

$$E\left[\widetilde{\nu}^h | Y^{h*} > 0\right] = \sigma^h \lambda^h \tag{7.70}$$

where λ^s and λ^h are defined as in equation (7.30) of the main article.

Furthermore, we assume a joint normal distribution of ε^s, ε^h, $\widetilde{\xi}_s^b$, $\widetilde{\xi}_s^s$, $\widetilde{\xi}_h^b$, and $\widetilde{\xi}_h^h$ with covariances $\sigma_s^{bs} = cov(\widetilde{\xi}_s^b, \varepsilon^s)$, $\sigma_s^{bh} = cov(\widetilde{\xi}_s^b, \varepsilon^h)$, $\sigma_s^{ss} = cov(\widetilde{\xi}_s^s, \varepsilon^s)$, $\sigma_s^{sh} = cov(\widetilde{\xi}_s^s, \varepsilon^h)$, $\sigma_h^{bs} = cov(\widetilde{\xi}_h^b, \varepsilon^s)$, $\sigma_h^{bh} = cov(\widetilde{\xi}_h^b, \varepsilon^h)$, $\sigma_h^{hs} = cov(\widetilde{\xi}_h^h, \varepsilon^s)$, and $\sigma_h^{hh} = cov(\widetilde{\xi}_h^h, \varepsilon^h)$, where $\widetilde{\xi}_s^b$, $\widetilde{\xi}_s^s$, $\widetilde{\xi}_h^b$, and $\widetilde{\xi}_h^h$ would be the error terms of equations (7.26), (7.27), (7.28), and (7.29) in the main article, respectively, without selectivity terms. From this we can obtain the conditional expectation of the error terms

$$E\left[\widetilde{\xi}_s^b | Y^{s*} > 0 \;\wedge\; Y^{h*} > 0\right] = \sigma_s^{bs}\,\lambda^{bs} + \sigma_s^{bh}\,\lambda^{bh} \tag{7.71}$$

$$E\left[\widetilde{\xi}_s^s | Y^{s*} > 0 \;\wedge\; Y^{h*} \leq 0\right] = \sigma_s^{ss}\,\lambda^{ss} + \sigma_s^{sh}\,\lambda^{sh} \tag{7.72}$$

$$E\left[\widetilde{\xi}_h^b | Y^{s*} > 0 \;\wedge\; Y^{h*} > 0\right] = \sigma_h^{bs}\,\lambda^{bs} + \sigma_h^{bh}\,\lambda^{bh} \tag{7.73}$$

$$E\left[\widetilde{\xi}_h^h | Y^{s*} \leq 0 \;\wedge\; Y^{h*} > 0\right] = \sigma_h^{hs}\,\lambda^{hs} + \sigma_h^{hh}\,\lambda^{hh} \tag{7.74}$$

where the λs are defined as in equations (7.31) to (7.33) of the main article.

7.A6 Proof of Selectivity Terms

In the following we derive the selectivity terms used in our 2SLS/IV estimation procedure.

To this end we consider a trivariate normal distribution of the variables X_1, X_2 and X_3 with density function $\phi_3(X_1, X_2, X_3)$, mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, where it holds:

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ & 1 & \rho \\ & & 1 \end{pmatrix} \tag{7.75}$$

The corresponding marginal normal distributions of the variables X_2 and X_3 are bivariate normal distributed with density function $\phi_2(X_1, X_2)$, mean vector $\boldsymbol{\mu}_{23}$ and covariance matrix $\boldsymbol{\Sigma}_{23}$, where it holds (see for example Greene, 2003):

$$\boldsymbol{\mu}_{23} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \qquad \boldsymbol{\Sigma}_{23} = \begin{pmatrix} 1 & \rho \\ & 1 \end{pmatrix} \tag{7.76}$$

The corresponding conditional distribution of X_1 has density function $\phi\left(X_1 \left|X_2, X_3\right.\right)$, mean μ_1^*, and variance σ_1^{2*}, where it holds (see for example Greene, 2003):

$$\mu_1^* = \frac{\left(\sigma_{12} - \rho\sigma_{13}\right) X_2 + \left(\sigma_{13} - \rho\sigma_{12}\right) X_3}{1-\rho^2} \tag{7.77}$$

$$\sigma_1^{2*} = \sigma_1^2 - \frac{\sigma_{12}^2 - 2\sigma_{12}\sigma_{13}\rho + \sigma_{13}^2}{1-\rho^2} \tag{7.78}$$

Given the definitions above we first prove the following three Lemmas

Lemma 1:

For a_2, $a_3 \in \mathbb{R}$ it holds

$$\int_{a_3}^{\infty} \phi\left(X_3\right) \phi\left(\frac{a_2 - \rho X_3}{\sqrt{1-\rho^2}}\right) \mathrm{d}X_3 = \sqrt{1-\rho^2}\phi\left(a_2\right) \Phi\left(\frac{-a_3 + \rho a_2}{\sqrt{1-\rho^2}}\right) \tag{7.79}$$

Proof:

$$\int_{a_3}^{\infty} \phi\left(X_3\right)\, \phi\left(\frac{a_2-\rho X_3}{\sqrt{1-\rho^2}}\right)\, \mathrm{d}X_3$$

$$= \int_{a_3}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}X_3^2}\, \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{a_2-\rho X_3}{\sqrt{1-\rho^2}}\right)^2}\, \mathrm{d}X_3 \tag{7.80}$$

$$= \int_{a_3}^{\infty} \frac{1}{2\pi}\, e^{-\frac{1}{2}\left(X_3^2+\frac{(a_2-\rho X_3)^2}{1-\rho^2}\right)}\, \mathrm{d}X_3 \tag{7.81}$$

$$= \int_{a_3}^{\infty} \frac{1}{2\pi}\, e^{-\frac{1}{2}\left(\frac{X_3^2\left(1-\rho^2\right)}{1-\rho^2}+\frac{a_2^2-2a_2\rho X_3+\rho^2 X_3^2}{1-\rho^2}\right)}\, \mathrm{d}X_3 \tag{7.82}$$

$$= \int_{a_3}^{\infty} \frac{1}{2\pi}\, e^{-\frac{1}{2}\left(\frac{a_2^2\left(1-\rho^2\right)}{1-\rho^2}+\frac{x_3^2-2x_3\rho a_2+\rho^2 a_2^2}{1-\rho^2}\right)}\, \mathrm{d}X_3 \tag{7.83}$$

$$= \int_{a_3}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}a_2^2}\, \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_3-\rho a_2}{\sqrt{1-\rho^2}}\right)^2}\, \mathrm{d}X_3 \tag{7.84}$$

$$= \int_{a_3}^{\infty} \phi\left(a_2\right)\, \phi\left(\frac{x_3-\rho a_2}{\sqrt{1-\rho^2}}\right)\, \mathrm{d}X_3 \tag{7.85}$$

$$= \phi\left(a_2\right)\sqrt{1-\rho^2}\int_{a_3}^{\infty} \frac{1}{\sqrt{1-\rho^2}}\phi\left(\frac{x_3-\rho a_2}{\sqrt{1-\rho^2}}\right)\, \mathrm{d}X_3 \tag{7.86}$$

$$= \phi\left(a_2\right)\sqrt{1-\rho^2}\int_{\frac{a_3-\rho a_2}{\sqrt{1-\rho^2}}}^{\infty} \phi\left(Z_3\right)\, \mathrm{d}Z_3 \tag{7.87}$$

$$= \sqrt{1-\rho^2}\phi\left(a_2\right)\, \Phi\left(\frac{-a_3+\rho a_2}{\sqrt{1-\rho^2}}\right) \tag{7.88}$$

q.e.d.

Corollary to Lemma 1:

$$\int_{-\infty}^{a_3} \phi\left(X_3\right)\phi\left(\frac{a_2-\rho X_3}{\sqrt{1-\rho^2}}\right)\, \mathrm{d}X_3 \;=\; \sqrt{1-\rho^2}\phi\left(a_2\right)\Phi\left(\frac{a_3-\rho a_2}{\sqrt{1-\rho^2}}\right) \tag{7.89}$$

Lemma 2

For a_2, $a_3 \in \mathbb{R}$ it holds

$$\begin{aligned}
& \int_{a_3}^{\infty} X_3 \, \phi\left(X_3\right) \; \Phi\left(\frac{-a_2+\rho X_3}{\sqrt{1-\rho^2}}\right) \, \mathrm{d}X_3 \\
= \; & \phi\left(a_3\right) \; \Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right) + \rho \, \phi\left(a_2\right) \; \Phi\left(\frac{-a_3+\rho a_2}{\sqrt{1-\rho^2}}\right) \qquad (7.90)
\end{aligned}$$

Proof:

$$\begin{aligned}
& \int_{a_3}^{\infty} X_3 \, \phi\left(X_3\right) \; \Phi\left(\frac{-a_2+\rho X_3}{\sqrt{1-\rho^2}}\right) \, \mathrm{d}X_3 \\
= \; & \int_{a_3}^{\infty} g'\left(X_3\right) \; f\left(X_3\right) \, \mathrm{d}X_3 \qquad (7.91)
\end{aligned}$$

with

$$g'\left(X_3\right) = X_3 \, \phi\left(X_3\right) \qquad (7.92)$$

$$f\left(X_3\right) = \Phi\left(\frac{-a_2+\rho X_3}{\sqrt{1-\rho^2}}\right) \qquad (7.93)$$

From partial integration it follows

$$\begin{aligned}
& \int_{a_3}^{\infty} g'\left(X_3\right) \; f\left(X_3\right) \, \mathrm{d}X_3 \\
= \; & \lim_{a \to \infty} g\left(a\right) \; f\left(a\right) - g\left(a_3\right) \; f\left(a_3\right) - \int_{a_3}^{\infty} g\left(X_3\right) \; f'\left(X_3\right) \, \mathrm{d}X_3 \qquad (7.94)
\end{aligned}$$

with

$$g\left(X_3\right) = -\phi\left(X_3\right) \qquad (7.95)$$

$$f'\left(X_3\right) = \phi\left(\frac{-a_2+\rho X_3}{\sqrt{1-\rho^2}}\right) \qquad (7.96)$$

substituting (7.95) and (7.96) into (7.94) we get

$$\begin{aligned}
&\lim_{a\to\infty} g\left(a\right)\ f\left(a\right)-g\left(a_3\right)\ f\left(a_3\right)-\int_{a_3}^{\infty} g\left(X_3\right)\ f'\left(X_3\right)\,\mathrm{d}X_3 \\
=\ & \phi\left(a_3\right)\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right)+\frac{\rho}{\sqrt{1-\rho^2}}\int_{a_3}^{\infty}\phi\left(X_3\right)\ \phi\left(\frac{-a_2+\rho X_3}{\sqrt{1-\rho^2}}\right)\,\mathrm{d}X_3 \quad (7.97)\\
=\ & \phi\left(a_3\right)\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right)+\frac{\rho}{\sqrt{1-\rho^2}}\int_{a_3}^{\infty}\phi\left(X_3\right)\ \phi\left(\frac{a_2-\rho X_3}{\sqrt{1-\rho^2}}\right)\,\mathrm{d}X_3 \quad (7.98)
\end{aligned}$$

applying Lemma 1 results in

$$\begin{aligned}
&\phi\left(a_3\right)\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right)+\frac{\rho}{\sqrt{1-\rho^2}}\int_{a_3}^{\infty}\phi\left(X_3\right)\ \phi\left(\frac{a_2-\rho X_3}{\sqrt{1-\rho^2}}\right)\,\mathrm{d}X_3 \\
=\ & \phi\left(a_3\right)\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right)+\rho\,\phi\left(a_2\right)\Phi\left(\frac{-a_3+\rho a_2}{\sqrt{1-\rho^2}}\right) \quad (7.99)
\end{aligned}$$

q.e.d.

Corollary to Lemma 2:

$$\begin{aligned}
&\int_{-\infty}^{a_3} X_3\,\phi\left(X_3\right)\ \Phi\left(\frac{-a_2+\rho X_3}{\sqrt{1-\rho^2}}\right)\,\mathrm{d}X_3 \\
=\ & -\phi\left(a_3\right)\ \Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right)+\rho\,\phi\left(a_2\right)\ \Phi\left(\frac{a_3-\rho a_2}{\sqrt{1-\rho^2}}\right) \quad (7.100)
\end{aligned}$$

Lemma 3

For $a_2,\ a_3 \in \mathbb{R}$ it holds:

$$\int_{-\infty}^{\infty}\int_{a_2}^{\infty}\int_{a_3}^{\infty}\phi_3\left(X_1,X_2,X_3\right)\,\mathrm{d}X_3\,\mathrm{d}X_2\,\mathrm{d}X_1=\Phi_2\left(-a_2,-a_3,\mathbf{\Sigma}_{23}\right) \quad (7.101)$$

Corollary to Lemma 3:

$$\int_{-\infty}^{\infty}\int_{a_2}^{\infty}\int_{-\infty}^{a_3}\phi_3\left(X_1,X_2,X_3\right)\,\mathrm{d}X_3\,\mathrm{d}X_2\,\mathrm{d}X_1=\Phi_2\left(-a_2,a_3,\left(1-\rho^2\right)\mathbf{\Sigma}_{23}^{-1}\right) \quad (7.102)$$

Lemma 4:

$$\int X_2\phi\left(X_2\right)\,\mathrm{d}X_2=-\phi\left(X_2\right) \quad (7.103)$$

Proof:

$$\frac{\partial \phi\left(X_2\right)}{\partial X_2} \quad = \quad -X_2\, \phi\left(X_2\right) \tag{7.104}$$

q.e.d.

Theorem

Given a trivariate normal distribution as defined above. Then it holds for any $a_2,\, a_3 \in \mathbb{R}$:

(i) $E\left(X_1 \,|X_2 > a_2 \wedge X_3 > a_3\right)$

$$= \quad \frac{\sigma_{13}\phi\left(a_3\right)\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right) + \sigma_{12}\phi\left(a_2\right)\Phi\left(\frac{-a_3+\rho a_2}{\sqrt{1-\rho^2}}\right)}{\Phi_2\left(-a_2, -a_3, \mathbf{\Sigma}_{23}\right)} \tag{7.105}$$

(ii) $E\left(X_1 \,|X_2 > a_2 \wedge X_3 < a_3\right)$

$$= \quad \frac{-\sigma_{13}\phi\left(a_3\right)\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right) + \sigma_{12}\phi\left(a_2\right)\Phi\left(\frac{a_3-\rho a_2}{\sqrt{1-\rho^2}}\right)}{\Phi_2\left(-a_2, a_3, \left(1-\rho^2\right)\mathbf{\Sigma}_{23}^{-1}\right)} \tag{7.106}$$

Proof of (i):
It holds the definition

$$E\left(X_1 \,|X_2 > a_2 \wedge X_3 > a_3\right) \quad = \quad \frac{\int_{-\infty}^{\infty}\int_{a_2}^{\infty}\int_{a_3}^{\infty} X_1\, \phi_3\left(X_1, X_2, X_3\right)\, \mathrm{d}X_3\, \mathrm{d}X_2\, \mathrm{d}X_1}{\int_{-\infty}^{\infty}\int_{a_2}^{\infty}\int_{a_3}^{\infty} \phi_3\left(X_1, X_2, X_3\right)\, \mathrm{d}X_3\, \mathrm{d}X_2\, \mathrm{d}X_1} \tag{7.107}$$

Applying Lemma 3 results in

$$E\left(X_1 \,|X_2 > a_2 \wedge X_3 > a_3\right) \quad = \quad \frac{\int_{-\infty}^{\infty}\int_{a_2}^{\infty}\int_{a_3}^{\infty} X_1\, \phi_3\left(X_1, X_2, X_3\right)\, \mathrm{d}X_3\, \mathrm{d}X_2\, \mathrm{d}X_1}{\Phi_2\left(-a_2, -a_3, \mathbf{\Sigma}_{23}\right)} \tag{7.108}$$

Now it holds for any trivariate normal distribution

$$
\begin{aligned}
& \int_{-\infty}^{\infty}\int_{a_2}^{\infty}\int_{a_3}^{\infty} X_1\,\phi_3\left(X_1,X_2,X_3\right)\,\mathrm{d}X_3\,\mathrm{d}X_2\,\mathrm{d}X_1 \\
= & \int_{a_2}^{\infty}\int_{a_3}^{\infty} \phi_2\left(X_2,X_3\right)\int_{-\infty}^{\infty} X_1\,\phi_3\left(X_1\left|X_2,X_3\right.\right)\,\mathrm{d}X_1\,\mathrm{d}X_3\,\mathrm{d}X_2 \qquad (7.109)\\
= & \int_{a_2}^{\infty}\int_{a_3}^{\infty} \phi_2\left(X_2,X_3\right)\mu_1^*\,\mathrm{d}X_3\,\mathrm{d}X_2 \qquad (7.110)\\
= & \int_{a_2}^{\infty}\int_{a_3}^{\infty} \phi_2\left(X_2,X_3\right)\frac{\left(\sigma_{12}-\rho\sigma_{13}\right)X_2+\left(\sigma_{13}-\sigma_{12}\rho\right)X_3}{1-\rho^2}\,\mathrm{d}X_3\,\mathrm{d}X_2 \qquad (7.111)\\
= & \int_{a_2}^{\infty}\int_{a_3}^{\infty} \phi\left(X_3\right)\phi\left(X_2\left|X_3\right.\right)\left(K_2X_2+K_3X_3\right)\,\mathrm{d}X_3\,\mathrm{d}X_2 \qquad (7.112)
\end{aligned}
$$

with

$$
\begin{aligned}
K_2 &= \frac{\sigma_{12}-\rho\sigma_{13}}{1-\rho^2} \qquad (7.113)\\
K_3 &= \frac{\sigma_{13}-\rho\sigma_{12}}{1-\rho^2} \qquad (7.114)
\end{aligned}
$$

Now it holds

$$
\begin{aligned}
& \int_{a_2}^{\infty}\int_{a_3}^{\infty}\phi\left(X_3\right)\frac{1}{\sqrt{1-\rho^2}}\,\phi\left(\frac{X_2-\rho X_3}{\sqrt{1-\rho^2}}\right)\left(K_2X_2+K_3X_3\right)\,\mathrm{d}X_3\,\mathrm{d}X_2 \\
= & \int_{a_3}^{\infty}\phi\left(X_3\right)\left(K_2\int_{a_2}^{\infty}X_2\frac{1}{\sqrt{1-\rho^2}}\,\phi\left(\frac{X_2-\rho X_3}{\sqrt{1-\rho^2}}\right)\,\mathrm{d}X_2\right. \qquad (7.115)\\
& \left.+K_3X_3\int_{a_2}^{\infty}\frac{1}{\sqrt{1-\rho^2}}\,\phi\left(\frac{X_2-\rho X_3}{\sqrt{1-\rho^2}}\right)\,\mathrm{d}X_2\right)\,\mathrm{d}X_3 \\
= & \int_{a_3}^{\infty}\phi\left(X_3\right)\left(K_2\sqrt{1-\rho^2}\int_{a_2}^{\infty}\frac{1}{\sqrt{1-\rho^2}}\frac{X_2-\rho X_3}{\sqrt{1-\rho^2}}\phi\left(\frac{X_2-\rho X_3}{\sqrt{1-\rho^2}}\right)\,\mathrm{d}X_2\right. \qquad (7.116)\\
& \left.+\left(K_2\rho+K_3\right)X_3\int_{a_2}^{\infty}\frac{1}{\sqrt{1-\rho^2}}\phi\left(\frac{X_2-\rho X_3}{\sqrt{1-\rho^2}}\right)\,\mathrm{d}X_2\right)\,\mathrm{d}X_3 \\
= & \int_{a_3}^{\infty}\phi\left(X_3\right)\left(K_2\sqrt{1-\rho^2}\int_{\frac{a_2-\rho X_3}{\sqrt{1-\rho^2}}}^{\infty}Z_2\phi\left(Z_2\right)\,\mathrm{d}Z_2\right. \qquad (7.117)\\
& \left.+\left(K_2\rho+K_3\right)X_3\int_{\frac{a_2-\rho X_3}{\sqrt{1-\rho^2}}}^{\infty}\phi\left(Z_2\right)\,\mathrm{d}Z_2\right)\,\mathrm{d}X_3
\end{aligned}
$$

applying Lemma 4 we get

$$\begin{aligned}
& \int_{a_3}^{\infty} \phi\left(X_3\right) K_2\sqrt{1-\rho^2}\phi\left(\frac{a_2-\rho X_3}{\sqrt{1-\rho^2}}\right)\,\mathrm{d}X_3 \\
& + \int_{a_3}^{\infty} \phi\left(X_3\right)\,\left(K_2\rho+K_3\right)\,X_3\Phi\left(\frac{-a_2+\rho X_3}{\sqrt{1-\rho^2}}\right) \\
= \;& K_2\sqrt{1-\rho^2}\int_{a_3}^{\infty} \phi\left(X_3\right)\phi\left(\frac{a_2-\rho X_3}{\sqrt{1-\rho^2}}\right)\,\mathrm{d}X_3 \qquad (7.118) \\
& + \left(K_2\rho+K_3\right)\int_{a_3}^{\infty} X_3\,\phi\left(X_3\right)\,\Phi\left(\frac{-a_2+\rho X_3}{\sqrt{1-\rho^2}}\right)\,\mathrm{d}X_3
\end{aligned}$$

applying Lemma 1 and Lemma 2 we get

$$\begin{aligned}
& K_2\sqrt{1-\rho^2}\,\sqrt{1-\rho^2}\,\phi\left(a_2\right)\,\Phi\left(\frac{-a_3+\rho a_2}{\sqrt{1-\rho^2}}\right) \\
& + \left(K_2\rho+K_3\right)\left(\phi\left(a_3\right)\,\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right)+\rho\,\phi\left(a_2\right)\,\Phi\left(\frac{-a_3+\rho a_2}{\sqrt{1-\rho^2}}\right)\right) \\
= \;& \left(K_2\left(1-\rho^2\right)+\left(K_2\rho+K_3\right)\rho\right)\phi\left(a_2\right)\Phi\left(\frac{-a_3+\rho a_2}{\sqrt{1-\rho^2}}\right) \qquad (7.119) \\
& + \left(K_2\rho+K_3\right)\phi\left(a_3\right)\,\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right) \\
= \;& \left(K_2+K_3\rho\right)\phi\left(a_2\right)\Phi\left(\frac{-a_3+\rho a_2}{\sqrt{1-\rho^2}}\right) \qquad (7.120) \\
& + \left(K_2\rho+K_3\right)\phi\left(a_3\right)\,\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right)
\end{aligned}$$

substituting (7.114) and (7.113) for K_2 and K_3

$$\begin{aligned}
& \left(\frac{\sigma_{12}-\rho\sigma_{13}}{1-\rho^2}+\frac{\sigma_{13}-\rho\sigma_{12}}{1-\rho^2}\rho\right)\phi\left(a_2\right)\Phi\left(\frac{-a_3+\rho a_2}{\sqrt{1-\rho^2}}\right) \\
& +\left(\frac{\sigma_{12}-\rho\sigma_{13}}{1-\rho^2}\rho+\frac{\sigma_{13}-\rho\sigma_{12}}{1-\rho^2}\right)\phi\left(a_3\right)\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right) \\
= & \left(\frac{\sigma_{12}-\rho\sigma_{13}+\sigma_{13}\rho-\rho^2\sigma_{12}}{1-\rho^2}\right)\phi\left(a_2\right)\Phi\left(\frac{-a_3+\rho a_2}{\sqrt{1-\rho^2}}\right) \qquad (7.121) \\
& +\left(\frac{\rho\sigma_{12}-\rho^2\sigma_{13}+\sigma_{13}-\rho\sigma_{12}}{1-\rho^2}\right)\phi\left(a_3\right)\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right) \\
= & \ \sigma_{12}\phi\left(a_2\right)\Phi\left(\frac{-a_3+\rho a_2}{\sqrt{1-\rho^2}}\right)+\sigma_{13}\phi\left(a_3\right)\Phi\left(\frac{-a_2+\rho a_3}{\sqrt{1-\rho^2}}\right) \qquad (7.122)
\end{aligned}$$

q.e.d.

Proof of (ii):
This proof is analogous to the proof of (i) except that the Corollaries are applied in place of the Lemmas.

7.A7 Formulas to Calculate Farm-Household Elasticities

7.A7.1 Notations

Price Elasticities on Production Side

$\varepsilon_{ij} = \dfrac{\partial X_i}{\partial P_j}\dfrac{P_j}{X_i}$ = traditional price elasticity of netput i with respect to price of netput j

$\varepsilon_{ij}^{FHM} = \dfrac{\partial X_i}{\partial P_j}\dfrac{P_j}{X_i}$ = FHM price elasticity of netput i with respect to price of netput/good j

Price Elasticities on Consumption Side

$\Theta_{ij} = \dfrac{\partial C_i}{\partial P_j}\dfrac{P_j}{C_i}$ = traditional Marshallian price elasticity of good i with respect to price of good j

$\Theta_{ij}^{H} = \dfrac{\partial C_i^H}{\partial P_j}\dfrac{P_j}{C_i}$ = traditional Hicksian price elasticity of good i with respect to price of good j

$\eta_i = \dfrac{\partial C_i}{\partial Y}\dfrac{Y}{C_i}$ = traditional income elasticity of good i

$\Theta_{ij}^{FHM} = \dfrac{\partial C_i}{\partial P_j}\dfrac{P_j}{C_i}$ = FHM price elasticity of good i with respect to price of netput/good j

Price Elasticities of Labor Allocation

$\varphi_{sL} = \dfrac{\partial X_L^s}{\partial P_L^s}\dfrac{P_L^s}{X_L^s}$ = traditional price elasticity of supplied labor with respect to labor price

$\varphi_{hL} = \dfrac{\partial X_L^h}{\partial P_L^h}\dfrac{P_L^h}{X_L^h}$ = traditional price elasticity of hired labor with respect to labor price

$\varphi_{sj}^{FHM} = \dfrac{\partial X_L^s}{\partial P_j}\dfrac{P_j}{X_L^s}$ = FHM price elasticity of supplied labor with respect to price of netput/good j

$\varphi_{hj}^{FHM} = \dfrac{\partial X_L^h}{\partial P_j}\dfrac{P_j}{X_L^h}$ = FHM price elasticity of hired labor with respect to price of netput/good j

$\varphi_{nj}^{FHM} = \dfrac{\partial X_L^n}{\partial P_j}\dfrac{P_j}{X_L^n}$ = FHM price elasticity of net supplied labor with respect to price of netput/good j

$\varphi_{fj}^{FHM} = \dfrac{\partial X_L^f}{\partial P_j}\dfrac{P_j}{X_L^f}$ = FHM price elasticity of family labor on the farm with respect to price of netput/good j

Shadow Price Elasticity of Labor

$\Psi_j = \dfrac{\partial P_L^*}{\partial P_j}\dfrac{P_j}{P_L^*}$ = elasticity of the shadow price of labor with respect to price of netput/good j

7.A7.2 Price Elasticities of the Separable Household Models

Price Elasticities on Production Side

The price elasticities on production side are simply the traditional price elasticities:

$$\varepsilon_{ij}^{FHM} = \varepsilon_{ij} \quad \forall\, i,j \in \{a,c,v,L\} \tag{7.123}$$

$$\varepsilon_{im}^{FHM} = 0 \quad \forall\, i \in \{a,c,v,L\} \tag{7.124}$$

Price Elasticities on Consumption Side

The price elasticities on consumption side consist of the normal Marshallian price effect and of an income effect due to an income change from farming or from working off-farm:

$$\Theta_{ij}^{sFHM} = \left.\frac{\partial C_i}{\partial P_j}\right|_{Y=\text{const.}} \frac{P_j}{C_i} + \frac{\partial C_i}{\partial Y}\frac{\partial Y}{\partial P_j}\frac{P_j}{C_i} \tag{7.125}$$

$$= \frac{\partial C_i^H}{\partial P_j}\frac{P_j}{C_i} + \frac{\partial C_i}{\partial Y}\frac{Y}{C_i}\left(\frac{\partial Y}{\partial P_j} - C_j\right)\frac{P_j}{Y} \tag{7.126}$$

$$= \Theta_{ij}^H + \eta_i\left(\frac{\partial Y}{\partial P_j} - C_j\right)\frac{P_j}{Y} \tag{7.127}$$

Evaluating $\frac{\partial Y}{\partial P_j}$ and removing all terms that are zero, we get the elasticities for each of the prices:

$$\Theta_{ij}^{sFHM} = \eta_i \frac{P_j X_j}{Y} \quad \forall\, i \in \{m, a, L\}\,,\; j \in \{c, v\} \tag{7.128}$$

$$\Theta_{ia}^{sFHM} = \Theta_{ia}^H + \eta_i \frac{P_a\left(X_a - C_a\right)}{Y} \quad \forall\, i \in \{m, a, L\} \tag{7.129}$$

$$\Theta_{iL}^{sFHM} = \Theta_{iL}^H + \eta_i \frac{P_j\left(X_L^s - X_L^h\right)}{Y} \quad \forall\, i \in \{m, a, L\} \tag{7.130}$$

$$\Theta_{im}^{sFHM} = \Theta_{im}^H - \eta_i \frac{P_m C_m}{Y} \quad \forall\, i \in \{m, a, L\} \tag{7.131}$$

Price Elasticity of Net Supply of Labor

The price elasticity of net supply of labor is calculated residually:

$$\varphi_{nj}^{sFHM} = \frac{\partial\left(X_L^s - X_L^h\right)}{\partial P_j}\frac{P_j}{X_L^{sn}} \tag{7.132}$$

$$= \frac{\partial\left(T_L + X_L - C_L\right)}{\partial P_j}\frac{P_j}{X_L^n} \tag{7.133}$$

$$= \frac{\partial X_L}{\partial P_j}\frac{P_j}{X_L}\frac{X_L}{X_L^n} - \frac{\partial C_L}{\partial P_j}\frac{P_j}{C_L}\frac{C_L}{X_L^n} \tag{7.134}$$

$$= \varepsilon_{Lj}^{FHM}\frac{X_L}{X_L^n} - \Theta_{Lj}^{FHM}\frac{C_L}{X_L^n} \quad \forall\, j \in \{a, c, v, L, m\} \tag{7.135}$$

7.A7.3 Price Elasticities of the Non-separable Household Models

The following formulas are valid for all four labor regimes. In case that the household does not supply labor, X_L^s and φ_L^s have to be set to zero and in case that the household does not hire labor, X_L^h and φ_L^h have to be set to zero.

Shadow Price Elasticities

We derive the shadow price elasticities from equation (7.14) of the main article:

$$\begin{aligned}
\Psi_j &= \frac{-\dfrac{\partial X_L}{\partial P_j} + \left.\dfrac{\partial C_L}{\partial P_j}\right|_{Y=\text{const.}} + \dfrac{\partial C_L}{\partial Y}\dfrac{\partial Y}{\partial P_j}}{\dfrac{\partial X_L}{\partial P_L^*} + \dfrac{\partial X_L^h}{\partial P_L^*} - \dfrac{\partial X_L^s}{\partial P_L^*} - \dfrac{\partial C_L^H}{\partial P_L^*}} \frac{P_j}{P_L^*} && (7.136)\\
&= \frac{-\dfrac{\partial X_L}{\partial P_j} + \dfrac{\partial C_L^H}{\partial P_j} + \dfrac{\partial C_L}{\partial Y}\left(\dfrac{\partial Y}{\partial P_j} - C_j\right)}{\dfrac{\partial X_L}{\partial P_L^*} + \dfrac{\partial X_L^h}{\partial P_L^*} - \dfrac{\partial X_L^s}{\partial P_L^*} - \dfrac{\partial C_L^H}{\partial P_L^*}} \frac{P_j}{P_L^*} && (7.137)\\
&= \frac{-\dfrac{\partial X_L}{\partial P_j}\dfrac{P_j}{X_L}X_L + \dfrac{\partial C_L^H}{\partial P_j}\dfrac{P_j}{C_L}C_L + \dfrac{\partial C_L}{\partial Y}\dfrac{Y}{C_L}\left(\dfrac{\partial Y}{\partial P_j} - C_j\right)\dfrac{P_j}{Y}C_L}{\dfrac{\partial X_L}{\partial P_L^*}\dfrac{P_L^*}{X_L}X_L + \dfrac{\partial X_L^h}{\partial P_L^*}\dfrac{P_L^*}{X_L^h}X_L^h - \dfrac{\partial X_L^s}{\partial P_L^*}\dfrac{P_L^*}{X_L^s}X_L^s - \dfrac{\partial C_L^H}{\partial P_L^*}\dfrac{P_L^*}{C_L}C_L} && (7.138)\\
&= \frac{-\varepsilon_{Lj}X_L + \Theta_{Lj}^H C_L + \eta_L\left(\dfrac{\partial Y}{\partial P_j} - C_j\right)\dfrac{P_j}{Y}C_L}{\varepsilon_{LL}X_L + \varphi_L^h X_L^h - \varphi_L^s X_L^s - \Theta_{LL}^H C_L} && (7.139)
\end{aligned}$$

Evaluating $\dfrac{\partial Y}{\partial P_j}$ and removing all terms that are zero, we get the elasticities for each of the exogenous prices:

$$\begin{aligned}
\Psi_j &= \frac{-\varepsilon_{Lj}X_L + \eta_L\dfrac{P_j X_j}{Y}C_L}{\varepsilon_{LL}X_L + \varphi_L^h X_L^h - \varphi_L^s X_L^s - \Theta_{LL}^H C_L}\ \forall j \in \{c, v\} && (7.140)\\
\Psi_a &= \frac{-\varepsilon_{La}X_L + \Theta_{La}^H C_L + \eta_L\dfrac{P_a\left(X_a - C_a\right)}{Y}C_L}{\varepsilon_{LL}X_L + \varphi_L^h X_L^h - \varphi_L^s X_L^s - \Theta_{LL}^H C_L} && (7.141)\\
\Psi_m &= \frac{\Theta_{Lm}^H C_L - \eta_L\dfrac{P_L C_L}{Y}C_L}{\varepsilon_{LL}X_L + \varphi_L^h X_L^h - \varphi_L^s X_L^s - \Theta_{LL}^H C_L} && (7.142)
\end{aligned}$$

Given the convexity of the profit function $\Pi(.)$ in netput prices and the concavity of the expenditure function $e(.)$ in commodity prices and assuming that $g(.)$ is convex in X_L^h and $f(.)$ is concave in X_L^s, the denominator is always positive, because $\varphi_L^h = \left(\partial^2 g/\partial {X_L^h}^2\right)^{-1}\left(P_L^h/X_L^h\right) \geq 0$, $X_L^h \geq 0$, $\varepsilon_{LL} = \left(\partial^2\Pi/\partial {P_L}^2\right)\left(P_L/X_L\right) \leq 0$, $X_L \leq 0$, $\varphi_L^s = \left(\partial^2 f/\partial {X_L^s}^2\right)^{-1}\left(P_L^s/X_L^s\right) \leq 0$, $X_L^s \geq 0$, $\Theta_{LL}^H = \left(\partial^2 e/\partial {P_L}^2\right)\left(C_L/P_L\right) \leq 0$, and $C_L \geq 0$.

Price Elasticities on Production Side

We derive the price elasticities on production side from equation (7.13) of the main article:

$$\begin{aligned}\varepsilon_{ij}^{iFHM} &= \left.\frac{\partial X_i}{\partial P_j}\right|_{P_L^*=\text{const.}} \frac{P_j}{X_i} + \frac{\partial X_i}{\partial P_L^*}\frac{P_L^*}{X_i}\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{P_L} && (7.143)\\ &= \varepsilon_{ij}^{sFHM} + \varepsilon_{iL}\Psi_j && (7.144)\end{aligned}$$

Substituting the direct component, which is the price elasticity of the separable model ε_{ij}^{sFHM}, we get the elasticities for each of the exogenous prices:

$$\begin{aligned}\varepsilon_{ij}^{iFHM} &= \varepsilon_{ij} + \varepsilon_{iL}\Psi_j \quad \forall\, i \in \{a, c, v, L\}\,,\; j \in \{c, a, v\} && (7.145)\\ \varepsilon_{im}^{iFHM} &= \varepsilon_{iL}\Psi_m \quad \forall\, i \in \{a, c, v, L\} && (7.146)\end{aligned}$$

Price Elasticities on Consumption Side

We derive the price elasticities on consumption side from equation (7.13) of the main article:

$$\begin{aligned}\Theta_{ij}^{iFHM} &= \left.\frac{\partial C_i}{\partial P_j}\right|_{P_L^*=\text{const.}} \frac{P_j}{C_i} + \frac{\partial C_i}{\partial P_L^*}\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{C_i} && (7.147)\\ &= \left.\frac{\partial C_i}{\partial P_j}\right|_{P_L^*=\text{const.}} \frac{P_j}{C_i} + \left(\left.\frac{\partial C_i}{\partial P_L^*}\right|_{Y=\text{const.}} + \frac{\partial C_i}{\partial Y}\frac{\partial Y}{\partial P_L^*}\right)\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{C_i} && (7.148)\\ &= \left.\frac{\partial C_i}{\partial P_j}\right|_{P_L^*=\text{const.}} \frac{P_j}{C_i} + \frac{\partial C_i^H}{\partial P_L^*}\frac{P_L^*}{C_i}\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{P_L^*} && (7.149)\\ &= \Theta_{ij}^{sFHM} + \Theta_{iL}^{H}\Psi_j && (7.150)\end{aligned}$$

Substituting the direct component, which is the price elasticity of the separable model Θ_{ij}^{sFHM}, we get the elasticities for each of the exogenous prices:

$$\begin{aligned}\Theta_{ij}^{iFHM} &= \eta_i\frac{P_jX_j}{Y} + \Theta_{iL}^{H}\Psi_j \quad \forall\, i \in \{m, a, L\}\,,\; j \in \{c, v\} && (7.151)\\ \Theta_{ia}^{iFHM} &= \Theta_{ia}^{H} + \eta_i\frac{P_a\left(X_a - C_a\right)}{Y} + \Theta_{iL}^{H}\Psi_a \quad \forall\, i \in \{m, a, L\} && (7.152)\\ \Theta_{im}^{iFHM} &= \Theta_{im}^{H} - \eta_i\frac{P_mC_m}{Y} + \Theta_{iL}^{H}\Psi_m \quad \forall\, i \in \{m, a, L\} && (7.153)\end{aligned}$$

Price Elasticities of Labor Allocation

We derive the price elasticities of labor supply and demand from equation (7.13) of the main article. Since the labor supply and demand do not directly depend on the exogenous prices, the direct component is zero:

$$\begin{aligned}
\varphi_{sj}^{iFHM} &= \frac{\partial X_L^s}{\partial P_L^*}\frac{P_L^*}{X_L^s}\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{P_L^*} && (7.154)\\
&= \varphi_L^s \Psi_j \quad \forall\, j \in \{c, a, v, m\} && (7.155)\\
\varphi_{hj}^{iFHM} &= \frac{\partial X_L^h}{\partial P_L^*}\frac{P_L^*}{X_L^h}\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{P_L^*} && (7.156)\\
&= \varphi_L^h \Psi_j \quad \forall\, j \in \{c, a, v, m\} && (7.157)
\end{aligned}$$

The remaining labor allocation elasticities are calculated residually:

$$\begin{aligned}
\varphi_{nj}^{iFHM} &= \frac{\partial \left(X_L^s - X_L^h\right)}{\partial P_L^*}\frac{P_L^*}{X_L^n}\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{P_L^*} && (7.158)\\
&= \frac{\partial X_L^s}{\partial P_L^*}\frac{P_L^*}{X_L^s}\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{P_L^*}\frac{X_L^s}{X_L^n} - \frac{\partial X_L^h}{\partial P_L^*}\frac{P_L^*}{X_L^h}\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{P_L^*}\frac{X_L^s}{X_L^n} && (7.159)\\
&= \varphi_j^s \frac{X_L^s}{X_L^n} - \varphi_j^h \frac{X_L^h}{X_L^n}\; j \in \{c, a, v, m\} && (7.160)\\
\varphi_{fj}^{iFHM} &= \frac{\partial \left(T_L - X_L^s - C_L^h\right)}{\partial P_L^*}\frac{P_L^*}{X_L^f}\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{P_L^*} && (7.161)\\
&= -\frac{\partial X_L^s}{\partial P_L^*}\frac{P_L^*}{X_L^s}\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{P_L^*}\frac{X_L^s}{X_L^f} - \frac{\partial C_L}{\partial P_L^*}\frac{P_L^*}{C_L}\frac{\partial P_L^*}{\partial P_j}\frac{P_j}{P_L^*}\frac{C_L}{X_L^f} && (7.162)\\
&= -\varphi_j^s \frac{X_L^s}{X_L^f} - \Theta_{Lj}^{iFHM}\frac{C_L}{X_L^f}\; j \in \{c, a, v, m\} && (7.163)
\end{aligned}$$

7.A8 Data Description

Table 7.A2: Characteristics of the Sample

Variable	Unit	Mean	Minimum	Maximum	Std.deviation
N_k	number	1.3	0.0	5.0	1.2
N_w	number	2.8	0.0	7.0	1.3
N_o	number	0.7	0.0	3.0	0.8
A_h	years	43	20	76	11
T_L	hours	11399	3650	27375	4457
$\|X_L\|$	hours	3686	400	9843	1717
X_L^h	hours	211	0	2085	365
X_L^s	hours	446	0	4000	876
X_L^n	hours	235	-2085	4000	1002
X_L^f	hours	3475	400	9236	1705
C_L	hours	7478	23	20873	4007
P_mC_m	1000 PLZ	91469	26365	280176	42853
P_aC_a	1000 PLZ	19041	1625	41853	7606
P_cX_c	1000 PLZ	132258	10451	1189412	133724
P_aX_a	1000 PLZ	212570	2669	2526524	239835
$P_v\|X_v\|$	1000 PLZ	211960	13480	2204671	213479
R_g	ha	14.7	1.2	101.5	12.4
R_k	1000 PLZ	649191	43960	4492025	554120
R_k/R_g	1000 PLZ / ha	46921	9170	215652	29039
N_c	number	0.9	0.0	3.0	0.6
W_u	%	19	9	25	4
W_i	km/100 km^2	58	39	71	9
W_t	1/1000 population	48	31	60	9
W_r	%	45	29	58	10
$\widetilde{P}_L$	Poland = 100	88	73	115	13
P_L^*	1000 PLZ/h	38	6	230	28

Note: Calculations are based on IERiGZ (1995). PLZ = Polish Zloty. Variables: N_k = number of family members up to 14 years, N_w = number of family members between 15 and 60 years, N_o = number of family members older than 60 years, A_h = age of the household head, T_L = total time available, $|X_L|$ = labor input on the farm, X_L^h = hired labor, X_L^s = supplied labor, X_L^n = net supplied labor, X_L^f = family labor input on the farm, C_L = leisure, P_mC_m = value of consumed market goods, P_aC_a = value of consumed self-produced goods, P_cX_c = value of produced crop products, P_aX_a = value of produced animal products, $P_v|X_v|$ = value of utilized variable inputs, R_g = amount of land of the farm, R_k = amount of capital of the farm, N_c = number of cars owned by the household, W_u = regional unemployment rate, W_i = regional density of the road and railroad network, W_t = regional density of telephones, W_r = proportion of the population that lives in rural areas, $\widetilde{P}_L$ = relative average regional wage level, $P_.^*$ = endogenous shadow price of labor.

Table 7.A3: Characteristics of the Different Labor Regimes

Variable	Unit	All	Sup. & Dem.	Only Sup.	Only Dem.	Autarkic
Number		199	57	47	61	34
N_k	number	1.3	1.5	1.3	1.4	0.7
N_w	number	2.8	2.8	3.2	2.4	3.0
N_o	number	0.7	0.6	0.6	0.8	0.7
A_h	years	43	41	44	43	45
T_L	hours	11399	11110	12891	10082	12185
$\|X_L\|$	hours	3686	3579	3372	4040	3668
X_L^h	hours	211	278	0	430	0
X_L^s	hours	446	515	1266	0	0
X_L^n	hours	235	237	1266	-430	0
X_L^f	hours	3475	3301	3372	3610	3668
C_L	hours	7478	7295	8254	6473	8517
$P_m C_m$	1000 PLZ	91469	105939	78012	97792	74467
$P_a C_a$	1000 PLZ	19041	18487	19245	19939	18076
$P_c X_c$	1000 PLZ	132258	157581	65883	180020	95869
$P_a X_a$	1000 PLZ	212570	220643	123997	300046	164531
$P_v\|X_v\|$	1000 PLZ	211960	232143	117552	299629	151343
R_g	ha	14.7	16.9	9.4	18.3	11.7
R_k	1000 PLZ	649191	788881	425398	816534	424132
R_k/R_g	1000 PLZ / ha	46921	49666	48516	48134	37938
N_c	number	0.9	1.0	0.8	0.9	0.8
W_u	%	19	20	19	18	20
W_i	km/100 km^2	58	55	60	60	57
W_t	1/1000 popul.	48	47	49	49	47
W_r	%	45	44	50	43	46
$\widetilde{P}_L$	Poland = 100	88	85	90	89	88
P_L^*	1000 PLZ/h	38	46	30	44	28

Note: Calculations are based on IERiGZ (1995). PLZ = Polish Zloty. Variables: N_k = number of family members up to 14 years, N_w = number of family members between 15 and 60 years, N_o = number of family members older than 60 years, A_h = age of the household head, T_L = total time available, $|X_L|$ = labor input on the farm, X_L^h = hired labor, X_L^s = supplied labor, X_L^n = net supplied labor, X_L^f = family labor input on the farm, C_L = leisure, $P_m C_m$ = value of consumed market goods, $P_a C_a$ = value of consumed self-produced goods, $P_c X_c$ = value of produced crop products, $P_a X_a$ = value of produced animal products, $P_v|X_v|$ = value of utilized variable inputs, R_g = amount of land of the farm, R_k = amount of capital of the farm, N_c = number of cars owned by the household, W_u = regional unemployment rate, W_i = regional density of the road and railroad network, W_t = regional density of telephones, W_r = proportion of the population that lives in rural areas, $\widetilde{P}_L$ = relative average regional wage level, $P_.^*$ = endogenous shadow price of labor.

7.A9 Estimation Results

7.A9.1 First-Stage Profit Function

Table 7.A4: Estimation Results of the Unrestricted 1st-Stage Profit Function

Parameter	$i = c$		$i = a$		$i = v$	
	Coef.	(t-val)	Coef.	(t-val)	Coef.	(t-val)
α_i	-1.72	(-0.73)	20.1	(4.31)	-17.4	(-5.14)
β_{ic}	-14.8	(-1.12)	19.8	(2.68)	-4.92	(-0.37)
β_{ia}	19.8	(2.68)	61.6	(5.76)	-81.4	(-8.04)
β_{iv}	-4.92	(-0.37)	-81.4	(-8.04)	86.3	(5.08)
δ_{ig}	6258	(11.37)	1002	(0.93)	-4306	(-5.37)
δ_{ik}	0.0829	(5.77)	0.209	(7.47)	-0.111	(-5.36)
γ_{gg}	-1157392	(-6.45)				
γ_{gk}	36.7	(7.59)				
γ_{kk}	$-1.26 \cdot 10^{-3}$	(-9.79)				
R^2	0.709		0.286		0.685	

Note: For definitions of the estimated coefficients see equation (7.20) of the main article, where the subscripts c, a, v, g, and k indicate crop products, animal products, variable inputs, land, and capital, respectively. The standard errors of the coefficients that have not been directly estimated are calculated with the formula of Klein (1953, p. 258). Monotonicity is fulfilled at 100% of the observations.

Table 7.A5: Estimation Results of the 1st-Stage Profit Function with Convexity Imposed

Parameter	$i = c$		$i = a$		$i = v$	
	Coef.	(t-val)	Coef.	(t-val)	Coef.	(t-val)
α_i	-2.28	(-0.57)	20.3	(3.16)	-17.0	(-3.21)
β_{ic}	3.31	(0.81)	14.6	(2.34)	-17.9	(-1.99)
β_{ia}	14.6	(2.34)	64.7	(2.93)	-79.3	(-3.16)
β_{iv}	-17.9	(-1.99)	-79.3	(-3.16)	97.3	(3.30)
δ_{ig}	6170	(4.60)	1024	(0.59)	-4294	(-2.26)
δ_{ik}	0.0855	(2.92)	0.208	(4.81)	-0.110	(-3.87)
γ_{gg}	-1149343	(-1.72)				
γ_{gk}	36.6	(1.89)				
γ_{kk}	$-1.26{\cdot}10^{-3}$	(-2.26)				
R^2	0.708		0.283		0.686	

Note: For definitions of the estimated coefficients see equation (7.20) of the main article, where the subscripts c, a, v, g, and k indicate crop products, animal products, variable inputs, land, and capital, respectively. The standard errors of the coefficients are calculated using the bootstrap resampling method (Efron, 1979; Efron and Tibshirani, 1993). Monotonicity is fulfilled at 100% of the observations. The R^2 values are almost identical to the model without convexity imposed, indicating that the data do not unreasonably contradict the convexity constraint (see table 7.A4).

Shadow Prices of Labor

One estimated shadow price is negative. The other shadow prices have a mean of 38498 PLZ/h and a median of 30236 PLZ/h. In 1994 the average gross wage in Poland was 32820 PLZ/h. 68% of the estimated shadow prices deviate less than 50% from this value.

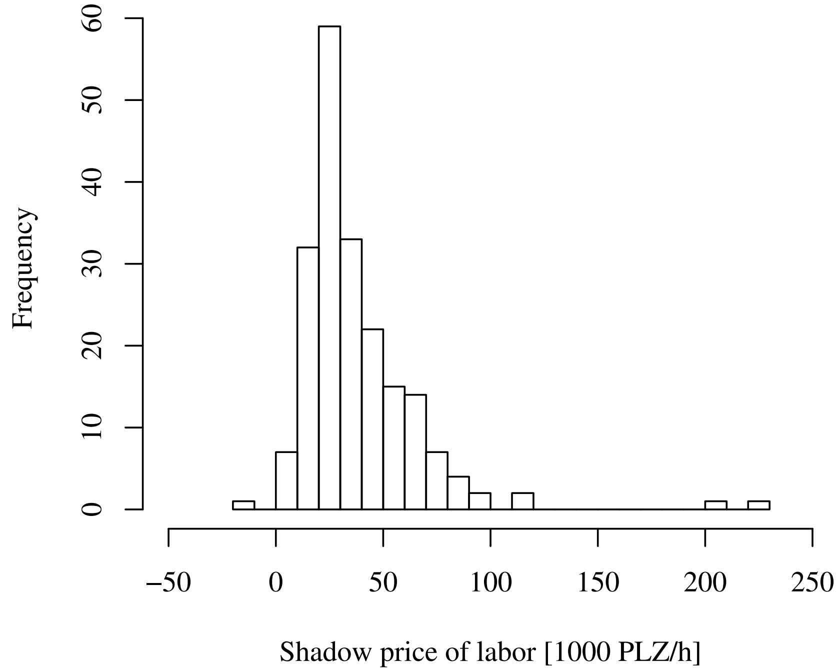


Figure 7.A1: Distribution of the estimated shadow prices of labor

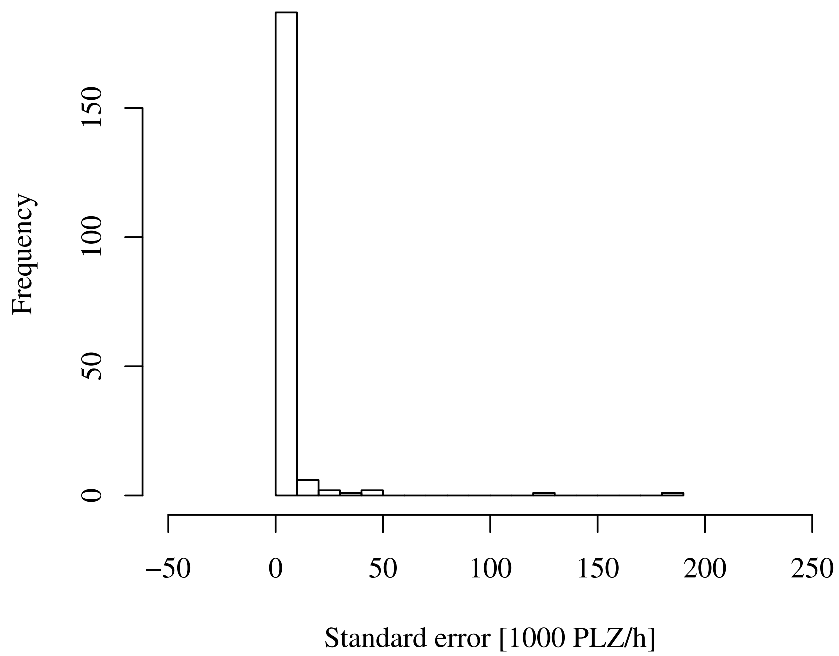
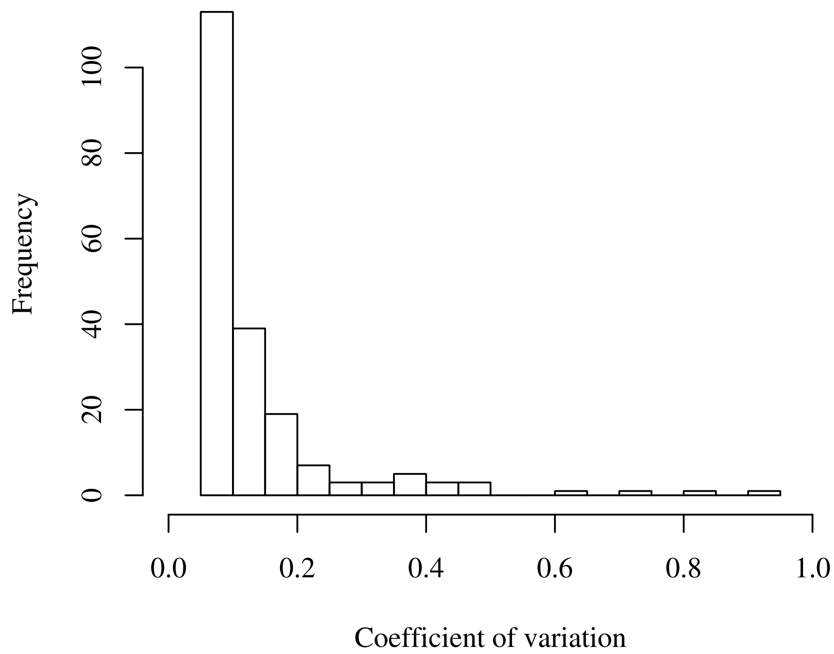


Figure 7.A2: Distribution of the standard errors of the estimated shadow prices of labor



Note: Only coefficients of variation of positive shadow prices are shown.

Figure 7.A3: Coefficients of variation of the estimated shadow prices of labor

7.A9.2 Second-Stage Profit Function

Table 7.A6: Estimation Results of the Unrestricted 2nd-Stage Profit Function

Parameter	$i=c$		$i=a$		$i=v$		$i=L$	
	Coef.	(t-val)	Coef.	(t-val)	Coef.	(t-val)	Coef.	(t-val)
α_i	-28774	(-3.22)	32491	(2.05)	-6714	(-0.57)	-62854	(-12.61)
β_{ic}	879	(0.02)	95377	(2.76)	-61671	(-1.14)	-34585	(-4.22)
β_{ia}	95377	(2.76)	76676	(1.19)	-162987	(-2.97)	-9066	(-0.63)
β_{iv}	-61671	(-1.14)	-162987	(-2.97)	221688	(2.95)	2970	(0.24)
β_{iL}	-34585	(-4.22)	-9066	(-0.63)	2970	(0.24)	40681	(7.48)
δ_{ig}	6896	(11.68)	131	(0.12)	-6000	(-7.02)	-3158	(-8.95)
δ_{ik}	0.121	(9.02)	0.292	(12.21)	-0.166	(-9.31)	$7.41{\cdot}10^{-3}$	(0.93)
γ_{gg}	-173	(-3.55)						
γ_{gk}	$9.88{\cdot}10^{-3}$	(9.24)						
γ_{kk}	$-3.55{\cdot}10^{-7}$	(-24.28)						
R^2	0.746		0.494		0.821		0.283	

Note: For definitions of the estimated coefficients see equation (7.15) of the main article, where the subscripts c, a, v, L, g, and k indicate crop products, animal products, variable inputs, labor, land, and capital, respectively. The standard errors of the coefficients that have not been directly estimated are calculated with the formula of Klein (1953, p. 258). Monotonicity is fulfilled at 98.0% of the observations. The estimation results with convexity imposed are presented in the main article, table 7.2.

Table 7.A7: Estimation Results of the 2nd-Stage Profit Function with Convexity Imposed

Parameter	$i = c$		$i = a$		$i = v$		$i = L$	
	Coef.	(t-val)	Coef.	(t-val)	Coef.	(t-val)	Coef.	(t-val)
α_i	-31261	(-2.31)	33699	(2.07)	-5480	(-0.37)	-62939	(-6.95)
β_{ic}	53083	(1.86)	64866	(2.75)	-84580	(-2.13)	-33368	(-3.46)
β_{ia}	64866	(2.75)	116773	(2.47)	-168328	(-2.68)	-13311	(-0.63)
β_{iv}	-84580	(-2.13)	-168328	(-2.68)	247344	(2.72)	5564	(0.32)
β_{iL}	-33368	(-3.46)	-13311	(-0.63)	5564	(0.32)	41115	(6.28)
δ_{ig}	6815	(4.59)	303	(0.14)	-6087	(-4.04)	-3181	(-2.81)
δ_{ik}	0.124	(4.40)	0.291	(7.49)	-0.167	(-6.97)	$7.87{\cdot}10^{-3}$	(0.20)
γ_{gg}	-172	(-1.28)						
γ_{gk}	$9.84{\cdot}10^{-3}$	(2.09)						
γ_{kk}	$-3.55{\cdot}10^{-7}$	(-2.26)						
R^2	0.747		0.492		0.821		0.278	

Note: For definitions of the estimated coefficients see equation (7.15) of the main article, where the subscripts c, a, v, L, g, and k indicate crop products, animal products, variable inputs, labor, land, and capital, respectively. The standard errors of the coefficients are calculated using the bootstrap resampling method (Efron, 1979; Efron and Tibshirani, 1993). Monotonicity is fulfilled at 97.0% of the observations. The R^2 values are almost identical to the model without convexity imposed, indicating that the data do not unreasonably contradict the convexity constraint (see table 7.A6).

Table 7.A8: Price Elasticities of the Restricted 2nd-Stage Profit Function

	P_c		P_a		P_v		P_L	
	Coef.	(t-val.)	Coef.	(t-val.)	Coef.	(t-val.)	Coef.	(t-val.)
X_c	0.429	(1.99)	0.503	(2.90)	-0.567	(-2.03)	-0.364	(-3.77)
X_a	0.320	(2.90)	0.533	(2.49)	-0.735	(-2.62)	-0.117	(-0.88)
X_v	0.356	(2.03)	0.726	(2.62)	-1.081	(-2.69)	-0.001	(-0.01)
X_L	0.340	(3.77)	0.172	(0.88)	-0.002	(-0.01)	-0.511	(-6.29)

7.A9.3 AIDS Model

Table 7.A9: Estimation Results of the AIDS

Parameter	$i = m$		$i = a$		$i = L$	
	Coef.	(t-val.)	Coef.	(t-val.)	Coef.	(t-val.)
α_i	0.555	(9.86)	0.185	(14.79)	0.260	(4.18)
β_i	-0.170	(-9.15)	-0.031	(-7.36)	0.201	(9.95)
γ_{im}	0.034	(1.28)	0.021	(0.79)	-0.055	(-5.34)
γ_{ia}	0.021	(0.79)	0.010	(0.35)	-0.031	(-9.36)
γ_{iL}	-0.055	(-5.34)	-0.031	(-9.36)	0.086	(7.97)
R^2	0.409		0.585		0.504	

Note: For definitions of the estimated coefficients see equation (7.16), where the subscripts m, a, and L indicate purchased market goods, self-produced goods, and leisure, respectively. The standard errors of the coefficients that have not been directly estimated are calculated with the formula of Klein (1953, p. 258). α_0 is set to 10.8, because this value gives the highest likelihood value of the AIDS Model. Monotonicity is fulfilled at 99.5% of the observations and concavity is fulfilled at 88.4% of the observations.

Table 7.A10: Price and Income Elasticities of the AIDS Model

	P_m		P_a		P_L	
	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)
Hicksian Price Elasticities						
C_m	-0.554	(-5.67)	0.144	(1.53)	0.409	(8.59)
C_a	0.648	(1.55)	-0.782	(-1.80)	0.134	(2.55)
C_L	0.176	(8.58)	0.014	(2.77)	-0.190	(-8.53)
Marshallian Price Elasticities						
C_m	-0.667	(-6.80)	0.119	(1.26)	0.149	(2.09)
C_a	0.503	(1.20)	-0.814	(-1.88)	-0.200	(-2.62)
C_L	-0.194	(-9.46)	-0.070	(-13.34)	-1.045	(-31.28)
Income Elasticities						
Y	0.399	(6.08)	0.511	(7.70)	1.308	(42.25)

7.A9.4 Labor Market Estimations

The analysis of labor supply and demand of the households is summarized in table 7.4 of the main article. The bivariate probit estimation shows that labor demand and supply decisions are not significantly correlated in the sample (ρ is not significantly different from zero). The probability that a household supplies off-farm labor increases significantly with the number of household members of working age (N_w) and with the rural nature of the region (W_r).

The probability that a household demands labor significantly depends on the capital endowment (R_k), the endowment of family labor (N_w, N_o), the age of the head of the household

(A_h, A_h^2), and the rural nature of the region (W_r). As expected, the probability increases with the capital endowment and decreases with the endowment of family labor. We also observe the expected signs for the age and squared age of the household head, i.e. we observe a u-shaped relation between age and the probability to hire on-farm labor with the lowest probability at the age of 44.4 years. Furthermore, the probability to hire labor decreases with the rural nature of the region.

The effective off-farm wage is significantly influenced by the proportion of supplied labor (X_L^s/T_L), the number of family members of working age (N_w), the age of the head of the household (A_h, A_h^2), and the rural nature of the region (W_r). Larger households and those in more rural areas receive a significantly lower effective off-farm wage. The coefficients of the age and squared age of the household head have the expected signs; i.e. we observe an inverse u-shaped relation between age and the effective off-farm wage with the highest wage at the age of 44.2 years. The estimated parameter of the inverse Mill's ratio is not significantly different from zero, indicating that there is no sample selection bias. If an average household (see table 7.A2 of the main article) increases the amount of supplied labor by 1%, the marginal revenue decreases by 0.075%. If this household doubles the amount of supplied labor from 446 to 892 hours per year, the marginal revenue decreases from 38498 to 35618 PLZ per hour.

The effective on-farm wage is significantly influenced by the amount of hired labor (X_L^h), the capital intensity on the farm (R_k/R_g), the regional unemployment rate (W_u), the regional density of the road and railroad network (W_i), and the rurality of the region (W_r). As expected, farms with a higher degree of mechanization pay higher wages because better skills are required on these farms. The negative impact of the rural nature and the positive impact of the road and railroad network on the effective on-farm wage might reflect heterogeneity of the average regional wages that is not captured in the regional data published by the statistical office $(\widetilde{P}_L)$. The positive effect of the regional unemployment rate is counter-intuitive. However, it might be correlated with some other regional variable not included in the analysis. In contrast to the labor supply side, the estimated parameter of the inverse Mill's ratio is significantly different from zero, indicating that an OLS estimation for labor-hiring households would be biased due to non-random sample selection. If an average household (see table 7.A2 of the main article) increases the amount of hired labor by 1%, the marginal cost increases by 0.259%. If this household doubles the amount of hired labor from 211 to 422 hours per year, the marginal cost increases from 38498 to 48467 PLZ per hour.

7.A10 Estimated Farm-Household Elasticities

7.A10.1 Elasticities for Different Labor Regimes

Table 7.A11: Price Elasticities of the Separable FHM (Calculated at Average Values of All Households)

	P_c		P_a		P_v		P_L		P_m	
	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)
X_c	0.43	(1.99)	0.50	(2.90)	-0.57	(-2.03)	-0.36	(-3.77)	0.00	
X_a	0.32	(2.90)	0.53	(2.49)	-0.73	(-2.62)	-0.12	(-0.88)	0.00	
X_v	0.36	(2.03)	0.73	(2.62)	-1.08	(-2.69)	-0.00	(-0.01)	0.00	
X_L	0.34	(3.77)	0.17	(0.88)	-0.00	(-0.01)	-0.51	(-6.29)	0.00	
C_m	0.13	(6.08)	0.33	(3.26)	-0.21	(-6.08)	0.45	(4.20)	-0.67	(-6.80)
C_a	0.17	(7.70)	-0.55	(-1.25)	-0.27	(-7.70)	0.18	(0.41)	0.50	(1.20)
C_L	0.43	(42.25)	0.61	(39.18)	-0.69	(-42.25)	-0.07	(-3.22)	-0.19	(-9.46)
X_L^n	-19.15	(-13.18)	-22.20	(-7.11)	22.00	(9.08)	10.30	(7.07)	6.16	(9.46)
X_L^f	0.34	(3.77)	0.17	(0.88)	-0.00	(-0.01)	-0.51	(-6.29)	0.00	
P_L^*	0.00		0.00		0.00		1.00		0.00	

Table 7.A12: Price Elasticities of the Non-separable FHM for Households that Both Supply and Hire Labor (Calculated at Average Values of All Households)

	P_c		P_a		P_v		P_m	
	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)
X_c	0.28	(1.51)	0.33	(2.06)	-0.39	(-1.53)	0.05	(2.67)
X_a	0.27	(2.40)	0.48	(2.30)	-0.68	(-2.26)	0.02	(0.87)
X_v	0.36	(2.10)	0.73	(2.57)	-1.08	(-2.61)	0.00	(0.01)
X_L	0.13	(1.43)	-0.08	(-0.50)	0.24	(1.98)	0.07	(3.32)
C_m	0.30	(6.21)	0.53	(4.76)	-0.41	(-6.74)	-0.72	(-7.32)
C_a	0.23	(7.53)	-0.48	(-1.13)	-0.33	(-8.25)	0.48	(1.16)
C_L	0.35	(15.54)	0.52	(18.27)	-0.60	(-21.22)	-0.17	(-7.98)
X_L^h	1.52	(1.46)	1.76	(1.26)	-1.75	(-1.37)	-0.49	(-1.26)
X_L^s	-6.26	(-3.79)	-7.25	(-3.56)	7.19	(3.69)	2.01	(3.55)
X_L^n	-13.25	(-3.46)	-15.37	(-3.42)	15.23	(3.45)	4.26	(3.42)
X_L^f	0.04	(0.16)	-0.19	(-0.60)	0.37	(1.22)	0.10	(1.28)
P_L^*	0.42	(3.94)	0.49	(3.68)	-0.48	(-3.82)	-0.13	(-3.67)

Table 7.A13: Price Elasticities of the Non-separable FHM for Households that Both Supply and Hire Labor (Calculated at Average Values of Households in this Labor Regime)

	P_c		P_a		P_v		P_m	
	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)
X_c	0.28	(1.53)	0.33	(2.09)	-0.40	(-1.55)	0.05	(2.56)
X_a	0.27	(2.44)	0.48	(2.31)	-0.68	(-2.27)	0.02	(0.86)
X_v	0.36	(2.10)	0.73	(2.57)	-1.08	(-2.62)	0.00	(0.01)
X_L	0.14	(1.51)	-0.06	(-0.41)	0.23	(1.84)	0.07	(3.11)
C_m	0.30	(5.87)	0.52	(4.64)	-0.40	(-6.42)	-0.72	(-7.28)
C_a	0.22	(7.50)	-0.49	(-1.13)	-0.33	(-8.20)	0.49	(1.17)
C_L	0.36	(15.23)	0.52	(17.96)	-0.60	(-20.81)	-0.17	(-7.99)
X_L^h	1.30	(1.35)	1.51	(1.18)	-1.50	(-1.27)	-0.42	(-1.18)
X_L^s	-5.47	(-3.46)	-6.34	(-3.28)	6.29	(3.38)	1.76	(3.28)
X_L^n	-13.41	(-3.12)	-15.55	(-3.10)	15.41	(3.11)	4.31	(3.10)
X_L^f	0.04	(0.14)	-0.19	(-0.55)	0.37	(1.12)	0.11	(1.17)
P_L^*	0.40	(3.60)	0.46	(3.39)	-0.46	(-3.51)	-0.13	(-3.40)

Note: To focus on the effect of the labor market regime, only X_L^s, X_L^h, z^s and z^h are the average values of households in this labor regime, while X_c, X_a, X_v, X_L, C_m, X_a and C_L are taken from the whole sample. $X_L^F = X_L - X_L^H$ and $T_L = X_L^S + X_L^F + C_L$ are calculated residually.

Table 7.A14: Price Elasticities of the Non-separable FHM for Households that only Supply Labor (Calculated at Average Values of Households in this Labor Regime)

	P_c		P_a		P_v		P_m	
	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)
X_c	0.24	(1.05)	0.29	(1.23)	-0.35	(-1.16)	0.06	(1.13)
X_a	0.26	(2.06)	0.46	(2.14)	-0.67	(-2.14)	0.02	(0.71)
X_v	0.36	(2.10)	0.73	(2.55)	-1.08	(-2.60)	0.00	(0.01)
X_L	0.08	(0.35)	-0.13	(-0.46)	0.30	(1.10)	0.08	(1.17)
C_m	0.34	(1.93)	0.57	(2.52)	-0.45	(-2.20)	-0.73	(-6.51)
C_a	0.24	(3.69)	-0.47	(-1.09)	-0.35	(-4.52)	0.48	(1.16)
C_L	0.34	(4.09)	0.50	(5.20)	-0.58	(-6.08)	-0.16	(-4.94)
X_L^s	-2.22	(-1.16)	-2.58	(-1.15)	2.55	(1.16)	0.72	(1.15)
X_L^f	0.08	(0.10)	-0.13	(-0.14)	0.30	(0.32)	0.08	(0.32)
P_L^*	0.51	(1.19)	0.59	(1.18)	-0.59	(-1.19)	-0.16	(-1.18)

Note: see note below table 7.A13.

Table 7.A15: Price Elasticities of the Non-separable FHM for Households that only Hire Labor (Calculated at Average Values of Households in this Labor Regime)

	P_c		P_a		P_v		P_m	
	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)
X_c	0.05	(0.33)	0.06	(0.41)	-0.13	(-0.56)	0.12	(3.32)
X_a	0.20	(1.28)	0.39	(1.65)	-0.59	(-1.69)	0.04	(0.91)
X_v	0.36	(1.97)	0.72	(2.37)	-1.08	(-2.46)	0.00	(0.01)
X_L	-0.19	(-2.56)	-0.45	(-4.59)	0.61	(6.94)	0.17	(5.35)
C_m	0.56	(9.20)	0.82	(7.10)	-0.70	(-9.84)	-0.80	(-8.25)
C_a	0.31	(5.71)	-0.38	(-0.92)	-0.43	(-6.70)	0.46	(1.11)
C_L	0.23	(8.30)	0.38	(10.99)	-0.46	(-13.92)	-0.13	(-6.62)
X_L^h	2.42	(0.40)	2.80	(0.40)	-2.78	(-0.40)	-0.78	(-0.40)
X_L^f	-0.54	(-8.30)	-0.88	(-10.99)	1.06	(13.92)	0.30	(6.62)
P_L^*	1.05	(8.01)	1.21	(7.39)	-1.20	(-7.92)	-0.34	(-5.61)

Note: see note below table 7.A13.

Table 7.A16: Price Elasticities of the Non-separable FHM for Autarkic Households (Calculated at Average Values of Households in this Labor Regime)

	P_c		P_a		P_v		P_m	
	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)	Elast.	(t-val.)
X_c	-0.07	(-0.51)	-0.07	(-0.49)	0.00	(0.01)	0.16	(3.40)
X_a	0.16	(0.88)	0.35	(1.32)	-0.55	(-1.44)	0.05	(0.92)
X_v	0.35	(1.84)	0.72	(2.24)	-1.08	(-2.37)	0.00	(0.01)
X_L	-0.35	(-5.77)	-0.63	(-9.15)	0.80	(11.88)	0.22	(6.00)
C_m	0.69	(10.50)	0.97	(8.51)	-0.85	(-11.88)	-0.85	(-8.75)
C_a	0.35	(5.16)	-0.34	(-0.82)	-0.48	(-6.07)	0.44	(1.08)
C_L	0.17	(5.77)	0.31	(9.15)	-0.39	(-11.88)	-0.11	(-6.00)
X_L^f	-0.35	(-5.77)	-0.63	(-9.15)	0.80	(11.88)	0.22	(6.00)
P_L^*	1.36	(9.17)	1.58	(9.44)	-1.56	(-9.79)	-0.44	(-5.65)

Note: see note below table 7.A13.

7.A10.2 Differences between Labor Regimes

Table 7.A17: Differences between Price Elasticities of the Separable FHM and the Non-separable FHM for Households that Supply as well as Demand Labor (Calculated at Average Values of All Households)

	P_c		P_a		P_v		P_m	
	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)
X_c	0.15	(2.43)	0.18	(2.73)	-0.18	(-2.55)	-0.05	(-2.67)
X_a	0.05	(0.87)	0.06	(0.78)	-0.06	(-0.81)	-0.02	(-0.87)
X_v	0.00	(0.01)	0.00	(0.01)	-0.00	(-0.01)	-0.00	(-0.01)
X_L	0.21	(3.37)	0.25	(2.85)	-0.25	(-3.08)	-0.07	(-3.32)
C_m	-0.17	(-3.64)	-0.20	(-3.42)	0.20	(3.54)	0.06	(4.02)
C_a	-0.06	(-2.16)	-0.07	(-2.09)	0.06	(2.14)	0.02	(2.19)
C_L	0.08	(3.64)	0.09	(3.41)	-0.09	(-3.54)	-0.03	(-4.00)
X_L^n	-5.90	(-1.50)	-6.84	(-1.38)	6.77	(1.44)	1.90	(1.47)
X_L^f	0.30	(1.21)	0.36	(1.24)	-0.37	(-1.28)	-0.10	(-1.28)
P_L^*	-0.42	(-3.94)	-0.49	(-3.68)	0.48	(3.82)	0.13	(3.67)

Table 7.A18: Differences between Price Elasticities of the Separable FHM and the Non-separable FHM for Households that Supply as well as Demand Labor

	P_c		P_a		P_v		P_m	
	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)
X_c	0.15	(2.34)	0.17	(2.60)	-0.17	(-2.45)	-0.05	(-2.56)
X_a	0.05	(0.87)	0.05	(0.78)	-0.05	(-0.80)	-0.02	(-0.86)
X_v	0.00	(0.01)	0.00	(0.01)	-0.00	(-0.01)	-0.00	(-0.01)
X_L	0.20	(3.14)	0.24	(2.71)	-0.23	(-2.90)	-0.07	(-3.11)
C_m	-0.16	(-3.37)	-0.19	(-3.19)	0.19	(3.29)	0.05	(3.66)
C_a	-0.05	(-2.10)	-0.06	(-2.03)	0.06	(2.08)	0.02	(2.13)
C_L	0.08	(3.37)	0.09	(3.18)	-0.09	(-3.29)	-0.02	(-3.65)
X_L^n	-5.74	(-1.30)	-6.66	(-1.21)	6.60	(1.26)	1.85	(1.28)
X_L^f	0.30	(1.08)	0.37	(1.13)	-0.38	(-1.17)	-0.11	(-1.17)
P_L^*	-0.40	(-3.60)	-0.46	(-3.39)	0.46	(3.51)	0.13	(3.40)

Table 7.A19: Differences between Price Elasticities of the Separable FHM and the Non-separable FHM for Households that only Supply Labor

	P_c		P_a		P_v		P_m	
	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)
X_c	0.19	(1.11)	0.22	(1.14)	-0.21	(-1.12)	-0.06	(-1.13)
X_a	0.06	(0.72)	0.07	(0.66)	-0.07	(-0.68)	-0.02	(-0.71)
X_v	0.00	(0.01)	0.00	(0.01)	-0.00	(-0.01)	-0.00	(-0.01)
X_L	0.26	(1.17)	0.30	(1.15)	-0.30	(-1.16)	-0.08	(-1.17)
C_m	-0.21	(-1.18)	-0.24	(-1.17)	0.24	(1.18)	0.07	(1.19)
C_a	-0.07	(-1.08)	-0.08	(-1.07)	0.08	(1.08)	0.02	(1.08)
C_L	0.10	(1.18)	0.11	(1.17)	-0.11	(-1.18)	-0.03	(-1.19)
X_L^n	-16.93	(-7.30)	-19.62	(-5.46)	19.45	(6.26)	5.45	(6.53)
X_L^f	0.26	(0.32)	0.30	(0.32)	-0.30	(-0.32)	-0.08	(-0.32)
P_L^*	-0.51	(-1.19)	-0.59	(-1.18)	0.59	(1.19)	0.16	(1.18)

Table 7.A20: Differences between Price Elasticities of the Separable FHM and the Non-separable FHM for Households that only Demand Labor

	P_c		P_a		P_v		P_m	
	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)
X_c	0.38	(2.97)	0.44	(3.73)	-0.44	(-3.26)	-0.12	(-3.32)
X_a	0.12	(0.92)	0.14	(0.81)	-0.14	(-0.84)	-0.04	(-0.91)
X_v	0.00	(0.01)	0.00	(0.01)	-0.00	(-0.01)	-0.00	(-0.01)
X_L	0.53	(6.19)	0.62	(4.30)	-0.61	(-5.01)	-0.17	(-5.35)
C_m	-0.43	(-6.86)	-0.50	(-6.38)	0.49	(6.75)	0.14	(8.39)
C_a	-0.14	(-2.52)	-0.16	(-2.45)	0.16	(2.51)	0.05	(2.52)
C_L	0.20	(6.88)	0.23	(6.35)	-0.23	(-6.76)	-0.06	(-8.27)
X_L^n	-21.57	(-3.43)	-25.00	(-3.23)	24.78	(3.34)	6.94	(3.35)
X_L^f	0.88	(9.79)	1.05	(6.42)	-1.06	(-7.87)	-0.30	(-6.62)
P_L^*	-1.05	(-8.01)	-1.21	(-7.39)	1.20	(7.92)	0.34	(5.61)

Table 7.A21: Differences between Price Elasticities of the Separable FHM and the Non-separable FHM for Autarkic Households

	P_c		P_a		P_v		P_m	
	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)
X_c	0.50	(3.06)	0.57	(4.04)	-0.57	(-3.43)	-0.16	(-3.40)
X_a	0.16	(0.93)	0.19	(0.82)	-0.18	(-0.85)	-0.05	(-0.92)
X_v	0.00	(0.01)	0.00	(0.01)	-0.00	(-0.01)	-0.00	(-0.01)
X_L	0.69	(7.97)	0.81	(4.99)	-0.80	(-6.00)	-0.22	(-6.00)
C_m	-0.56	(-8.19)	-0.65	(-8.22)	0.64	(8.52)	0.18	(9.49)
C_a	-0.18	(-2.58)	-0.21	(-2.54)	0.21	(2.59)	0.06	(2.55)
C_L	0.26	(8.24)	0.30	(8.17)	-0.30	(-8.57)	-0.08	(-9.36)
X_L^n	-19.15	(-13.18)	-22.20	(-7.11)	22.00	(9.08)	6.16	(9.46)
X_L^f	0.69	(7.97)	0.81	(4.99)	-0.80	(-6.00)	-0.22	(-6.00)
P_L^*	-1.36	(-9.17)	-1.58	(-9.44)	1.56	(9.79)	0.44	(5.65)

Table 7.A22: Differences between Price Elasticities of the Households that Supply as well as Demand Labor and the Households that only Supply Labor

	P_c		P_a		P_v		P_m	
	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)
X_c	0.04	(0.34)	0.05	(0.34)	-0.05	(-0.34)	-0.01	(-0.34)
X_a	0.01	(0.32)	0.01	(0.32)	-0.01	(-0.32)	-0.00	(-0.32)
X_v	0.00	(0.01)	0.00	(0.01)	-0.00	(-0.01)	-0.00	(-0.01)
X_L	0.06	(0.34)	0.07	(0.34)	-0.06	(-0.34)	-0.02	(-0.34)
C_m	-0.05	(-0.34)	-0.05	(-0.34)	0.05	(0.34)	0.01	(0.34)
C_a	-0.01	(-0.34)	-0.02	(-0.34)	0.02	(0.34)	0.00	(0.34)
C_L	0.02	(0.34)	0.02	(0.34)	-0.02	(-0.34)	-0.01	(-0.34)
X_L^h	1.30	(0.47)	1.51	(0.47)	-1.50	(-0.47)	-0.42	(-0.47)
X_L^s	-3.25	(-6.98)	-3.77	(-5.67)	3.73	(6.32)	1.05	(5.84)
X_L^n	-11.19	(-4.70)	-12.97	(-4.65)	12.85	(4.68)	3.60	(4.66)
X_L^f	-0.04	(-0.07)	-0.06	(-0.10)	0.08	(0.12)	0.02	(0.12)
P_L^*	-0.11	(-0.34)	-0.13	(-0.34)	0.13	(0.34)	0.04	(0.34)

Table 7.A23: Differences between Price Elasticities of the Households that Supply as well as Demand Labor and the Households that only Demand Labor

	P_c		P_a		P_v		P_m	
	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)
X_c	0.24	(2.59)	0.27	(3.12)	-0.27	(-2.81)	-0.08	(-2.77)
X_a	0.08	(0.92)	0.09	(0.81)	-0.09	(-0.84)	-0.02	(-0.91)
X_v	0.00	(0.01)	0.00	(0.01)	-0.00	(-0.01)	-0.00	(-0.01)
X_L	0.33	(4.17)	0.38	(3.52)	-0.38	(-3.81)	-0.11	(-3.77)
C_m	-0.26	(-4.17)	-0.31	(-4.20)	0.30	(4.23)	0.09	(4.28)
C_a	-0.09	(-2.28)	-0.10	(-2.25)	0.10	(2.29)	0.03	(2.25)
C_L	0.12	(4.17)	0.14	(4.20)	-0.14	(-4.23)	-0.04	(-4.27)
X_L^h	-1.11	(-0.21)	-1.29	(-0.21)	1.28	(0.21)	0.36	(0.21)
X_L^s	-5.47	(-3.46)	-6.34	(-3.28)	6.29	(3.38)	1.76	(3.28)
X_L^n	-15.83	(-1.53)	-18.35	(-1.53)	18.18	(1.53)	5.09	(1.53)
X_L^f	0.58	(2.13)	0.68	(2.11)	-0.69	(-2.17)	-0.19	(-2.11)
P_L^*	-0.65	(-4.25)	-0.75	(-4.31)	0.74	(4.33)	0.21	(3.64)

Table 7.A24: Differences between Price Elasticities of the Households that Supply as well as Demand Labor and the Autarkic Households

	P_c		P_a		P_v		P_m	
	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)
X_c	0.35	(2.88)	0.41	(3.76)	-0.40	(-3.22)	-0.11	(-3.09)
X_a	0.11	(0.94)	0.13	(0.83)	-0.13	(-0.86)	-0.04	(-0.93)
X_v	0.00	(0.01)	0.00	(0.01)	-0.00	(-0.01)	-0.00	(-0.01)
X_L	0.49	(6.25)	0.57	(4.68)	-0.56	(-5.32)	-0.16	(-4.89)
C_m	-0.39	(-5.79)	-0.46	(-6.27)	0.45	(6.15)	0.13	(5.76)
C_a	-0.13	(-2.47)	-0.15	(-2.47)	0.15	(2.50)	0.04	(2.42)
C_L	0.18	(5.82)	0.21	(6.26)	-0.21	(-6.19)	-0.06	(-5.74)
X_L^h	1.30	(1.35)	1.51	(1.18)	-1.50	(-1.27)	-0.42	(-1.18)
X_L^s	-5.47	(-3.46)	-6.34	(-3.28)	6.29	(3.38)	1.76	(3.28)
X_L^n	-13.41	(-3.12)	-15.55	(-3.10)	15.41	(3.11)	4.31	(3.10)
X_L^f	0.40	(1.41)	0.44	(1.32)	-0.42	(-1.30)	-0.12	(-1.30)
P_L^*	-0.96	(-5.76)	-1.11	(-6.30)	1.10	(6.15)	0.31	(4.32)

Table 7.A25: Differences between Price Elasticities of the Households that only Supply Labor and the Households that only Demand Labor

	P_c		P_a		P_v		P_m	
	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)
X_c	0.20	(1.14)	0.23	(1.18)	-0.22	(-1.16)	-0.06	(-1.16)
X_a	0.06	(0.75)	0.07	(0.69)	-0.07	(-0.71)	-0.02	(-0.74)
X_v	0.00	(0.01)	0.00	(0.01)	-0.00	(-0.01)	-0.00	(-0.01)
X_L	0.27	(1.22)	0.32	(1.20)	-0.31	(-1.21)	-0.09	(-1.21)
C_m	-0.22	(-1.22)	-0.25	(-1.22)	0.25	(1.22)	0.07	(1.22)
C_a	-0.07	(-1.11)	-0.08	(-1.11)	0.08	(1.11)	0.02	(1.11)
C_L	0.10	(1.22)	0.12	(1.22)	-0.12	(-1.22)	-0.03	(-1.22)
X_L^h	-2.42	(-0.30)	-2.80	(-0.30)	2.78	(0.30)	0.78	(0.30)
X_L^s	-2.22	(-1.16)	-2.58	(-1.15)	2.55	(1.16)	0.72	(1.15)
X_L^n	-4.64	(-0.58)	-5.38	(-0.58)	5.33	(0.58)	1.49	(0.58)
X_L^f	0.62	(0.76)	0.75	(0.79)	-0.76	(-0.82)	-0.21	(-0.81)
P_L^*	-0.54	(-1.22)	-0.62	(-1.22)	0.62	(1.22)	0.17	(1.20)

Table 7.A26: Differences between Price Elasticities of the Households that only Supply Labor and the Autarkic Households

	P_c		P_a		P_v		P_m	
	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)
X_c	0.31	(1.67)	0.36	(1.80)	-0.36	(-1.73)	-0.10	(-1.70)
X_a	0.10	(0.86)	0.12	(0.77)	-0.11	(-0.80)	-0.03	(-0.85)
X_v	0.00	(0.01)	0.00	(0.01)	-0.00	(-0.01)	-0.00	(-0.01)
X_L	0.43	(1.94)	0.50	(1.88)	-0.50	(-1.91)	-0.14	(-1.88)
C_m	-0.35	(-1.92)	-0.40	(-1.94)	0.40	(1.94)	0.11	(1.92)
C_a	-0.11	(-1.57)	-0.13	(-1.58)	0.13	(1.58)	0.04	(1.56)
C_L	0.16	(1.93)	0.19	(1.94)	-0.19	(-1.94)	-0.05	(-1.92)
X_L^s	-2.22	(-1.16)	-2.58	(-1.15)	2.55	(1.16)	0.72	(1.15)
X_L^n	-2.22	(-1.16)	-2.58	(-1.15)	2.55	(1.16)	0.72	(1.15)
X_L^f	0.43	(0.53)	0.50	(0.53)	-0.50	(-0.53)	-0.14	(-0.53)
P_L^*	-0.85	(-1.92)	-0.99	(-1.94)	0.98	(1.94)	0.27	(1.84)

Table 7.A27: Differences between Price Elasticities of the Households that only Demand Labor and the Autarkic Households

	P_c		P_a		P_v		P_m	
	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)	Diff.	(t-val.)
X_c	0.11	(2.28)	0.13	(2.72)	-0.13	(-2.47)	-0.04	(-2.34)
X_a	0.04	(0.94)	0.04	(0.83)	-0.04	(-0.86)	-0.01	(-0.92)
X_v	0.00	(0.01)	0.00	(0.01)	-0.00	(-0.01)	-0.00	(-0.01)
X_L	0.16	(3.32)	0.19	(3.10)	-0.18	(-3.22)	-0.05	(-2.97)
C_m	-0.13	(-3.11)	-0.15	(-3.30)	0.15	(3.23)	0.04	(3.01)
C_a	-0.04	(-2.07)	-0.05	(-2.09)	0.05	(2.10)	0.01	(2.01)
C_L	0.06	(3.12)	0.07	(3.30)	-0.07	(-3.23)	-0.02	(-3.01)
X_L^h	2.42	(0.40)	2.80	(0.40)	-2.78	(-0.40)	-0.78	(-0.40)
X_L^n	2.42	(0.40)	2.80	(0.40)	-2.78	(-0.40)	-0.78	(-0.40)
X_L^f	-0.18	(-4.46)	-0.24	(-5.22)	0.26	(5.63)	0.07	(4.74)
P_L^*	-0.31	(-3.03)	-0.36	(-3.21)	0.36	(3.14)	0.10	(2.68)

Bibliography

Abdulai, A., and C.L. Delgado. 1999. "Determinants of Nonfarm Earnings of Farm-Based Husbands and Wives in Northern Ghana." *American Journal of Agricultural Economics* 81:117–130.

Abdulai, A., P. Monnin, and J. Gerber. 2005. "Output Uncertainty, Information Acquisition and the Adoption of New Technologies." Unpublished Working Paper, University of Kiel, Germany.

Benjamin, D. 1992. "Household Composition, Labor Market, and Labor Demand: Testing for Separation in Agricultural Household Models." *Econometrica* 60:287–322.

Blundell, R.W., and J.M. Robin. 1999. "Estimation in Large and Disaggregated Demand Systems: An Estimator for Conditionally Linear Systems." *Journal of Applied Econometrics* 14:209–232.

Carter, M.R., and P. Olinto. 2003. "Getting Institutions "Right" for Whom? Credit Constraints and the Impact of Property Rights on the Quantity and Composition of Investment." *American Journal of Agricultural Economics* 85:173–186.

Carter, M.R., and Y. Yao. 2002. "Local versus Global Separability in Agricultural Household Models: The Factor Price Equalization Effect of Land Transfer Rights." *American Journal of Agricultural Economics* 84:702–715.

Chambers, R.G., and R.E. Lopez. 1987. "Tax Policies and the Financially Constrained Farm Household." *American Journal of Agricultural Economics* 69:369–377.

de Janvry, A., M. Fafchamps, and E. Sadoulet. 1991. "Peasant Household Behaviour with Missing Markets: Some Paradoxes Explained." *The Economic Journal* 101:1400–1417.

de Janvry, A., E. Sadoulet, M. Fafchamps, and M. Raki. 1992. "Structural Adjustment and the Peasantry in Morocco: A Computable Household Model." *European Review of Agricultural Economics* 19:427–453.

Deaton, A., and J. Muellbauer. 1980. "An Almost Ideal Demand System." *The American Economic Review* 70:312–326.

Deolalikar, A.B., and W.P.M. Vijverberg. 1983. "The Heterogeneity of Family and Hired Labour in Agricultural Production: A Test Using District-level Data from India." *Journal of Economic Development* 8:45–69.

—. 1987. "A Test of Heterogeneity of Family and Hired Labour in Asian Agriculture." *Oxford Bulletin of Economics and Statistics* 49:291–305.

Diewert, W.E., and T.J. Wales. 1987. "Flexible Functional Forms and Global Curvature Conditions." *Econometrica* 55:43–68.

—. 1992. "Quadratic Spline Models for Producer's Supply and Demand Functions." *International Economic Review* 33:705–722.

Efron, B. 1979. "Bootstrap Methods: Another Look at the Jackknife." *The Annals of Statistics* 7:1–26.

Efron, B., and R.J. Tibshirani. 1993. *An Introduction to the Bootstrap*, vol. 57 of *Monographs on Statistics and Applied Probability*. New York, London: Chapman & Hall.

Eswaran, M., and A. Kotwal. 1986. "Access to Capital and Agrarian Production Organization." *The Economic Journal* 96:482–498.

Fafchamps, M. 1992. "Cash Crop Production, Food Price Volatility, and Rural Market Integration in the Third World." *American Journal of Agricultural Economics* 74:90–99.

Feder, G. 1985. "The Relation between Farm Size and Farm Productivity: The Role of Family Labor, Supervision and Credit Constraints." *Journal of Development Economics* 18:297–313.

Finkelshtain, I., and J.A. Chalfant. 1991. "Marketed Surplus under Risk: Do Peasants Agree with Sandmo?" *American Journal of Agricultural Economics* 73:557–567.

Frisvold, G.B. 1994. "Does Supervision Matter? Some Hypothesis Tests Using Indian Farm-Level Data." *Journal of Development Economics* 43:217–238.

Genz, A., F. Bretz, and T. Hothorn. 2005. *mvtnorm: Multivariate Normal and t Distribution*. R package version 0.7-2, http://cran.r-project.org.

Genz, A., M. Meyer, T. Lumley, and M. Maechler. 2005. *adapt: Multidimensional Numerical Integration*. R package version 1.0-3, http://cran.r-project.org.

Glauben, T., C.H.C.A. Henning, and A. Henningsen. 2003. "Farm household decisions under various tax policies: Comparative static results and evidence from household data." Paper presented at the 25th International Conference of the IAAE, Durban, South Africa.

Glowny Urzad Statystyczny. 1996. "Informacja o Sytuacji Spoleczno - Gospodarczej Wojewodztw." Kwartalnik Nr 1-2, Warsaw.

Goetz, S.J. 1992. "A Selectivity Model of Household Food Marketing Behavior in Sub-Sahara Africa." *American Journal of Agricultural Economics* 74:444–452.

Gourieroux, C., A. Monfort, and A. Trognon. 1985. "Moindres Carres Asymptotiques." *Annales de l'INSEE* 58:91–122.

Granovetter, M.S. 1973. "The strength of weak ties." *The American Journal of Sociology* 78:1360–1380.

Greene, W.H. 2003. *Econometric Analysis*, 5th ed. Prentice Hall.

Hamann, J.D., and A. Henningsen. 2005. *systemfit: Simultaneous Equation Estimation*. R package version 0.7-6, http://cran.r-project.org.

Henning, C.H.C.A. 1994. *Unternehmens-Haushalts-Modelle*. Duncker & Humblot, Berlin.

Henning, C.H.C.A., and A. Henningsen. forthcoming. "AJAE Appendix: Modeling Price Response of Farm Households in Imperfect Labour Markets in Poland: Incorporating Transaction Costs and Heterogeneity into a Farm Household Approach." Unpublished manuscript. Available at: http://agecon.lib.umn.edu/.

Henningsen, A., and O. Toomet. 2005. *micEcon: Tools for Microeconomic Analysis and Microeconomic Modeling*. R package version 0.2-1, http://cran.r-project.org.

IERiGZ. 1995. "Indywiualne wyniki rachunkowosci rolnej gospodarstw chlopskich - Rok 1991-1994 - Makroregion VIII - Srodkowo Zachodni." Unpublished data, personally requested from IERiGZ, Warsaw, Poland.

Jacoby, H. 1993. "Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra." *Review of Economic Studies* 60:903–921.

Key, N., E. Sadoulet, and A. de Janvry. 2000. "Transaction Costs and Agricultural Household Supply Response." *American Journal of Agricultural Economics* 82:245–259.

Klein, L.R. 1953. *A Textbook of Econometrics*. New York: Row, Peterson and Co.

Kodde, D.A., F.C. Palm, and G.A. Pfann. 1990. "Asymptotic Least-Squares Estimation Efficiency Considerations and Applications." *Journal of Applied Econometrics* 5:229–243.

Koebel, B. 1998. "Tests of Representative Firm Models: Results for German Manufacturing Industries." *Journal of Productivity Analysis* 10:251–270.

Koebel, B., M. Falk, and F. Laisney. 2003. "Imposing and Testing Curvature Conditions on a Box-Cox Cost Function." *Journal of Business and Economic Statistics* 21:319–335.

Kohli, U.R. 1993. "A Symmetric Normalized Quadratic GNP Function and the US Demand for Imports and Supply of Exports." *International Economic Review* 34:243–255.

Lau, L.J. 1978a. "Applications of Profit Functions." In M. Fuss and D. McFadden, eds. *Production Economics: A Dual Approach to Theory and Applications, Vol 1*. Amsterdam: North-Holland, pp. 133–216.

—. 1978b. "Testing and Imposing Monotonicity, Convexity and Quasi-Convexity Constraints." In M. Fuss and D. McFadden, eds. *Production Economics: A Dual Approach to Theory and Applications, Vol. 1*. Amsterdam: North-Holland, pp. 409–453.

Lee, L., G.S. Maddala, and R.P. Trost. 1980. "Asymptotic Covariance Matrices of Two-stage Probit and Two-stage Tobit Methods for Simultaneous Equations Models with Selectivity." *Econometrica* 48:491–503.

Lopez, R. 1984. "Estimating Labor Supply and Production Decisions of Self-Employed Farm Producers." *European Economic Review* 24:61–82.

—. 1986. "Structural Models of the Farm Household that Allow for Interdependent Utility and Profit Maximization Decisions." In I. Singh, L. Squire, and J. Strauss, eds. *Agricultural Household Models*. Baltimore: The John Hopkins University Press, pp. 306–325.

Omamo, S.W. 1998. "Transportation Costs and Smallholder Cropping Choices: An Application to Siaya District, Kenya." *American Journal of Agricultural Economics* 80:1083–1095.

Pagan, A. 1984. "Econometric Issues in the Analysis of Regressions with Generated Regressors." *International Economic Review* 25:221–247.

R Development Core Team. 2005. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Sadoulet, E., A. de Janvry, and C. Benjamin. 1998. "Household Behavior with Imperfect Labor Markets." *Industrial Relations* 37:85–108.

—. 1996. "Labor Market Imperfections and Selective Separability in Household Models: A Predictive Typology of Mexican Ejidatarios." Working Paper No. 786, Department of Agricultural and Resource Economics, University of California at Berkeley.

Saha, A., H.A. Love, and R.B. Schwart. 1994. "Adoption of Emerging Technologies under Output Uncertainty." *American Journal of Agricultural Economics* 76:836–846.

Skoufias, E. 1994. "Using Shadow Wages to Estimate Labor Supply of Agricultural Households." *American Journal of Agricultural Economics* 76:215–227.

Spence, A.M. 1976. "Consumer Misperceptions, Product Failure and Producer Liability." *Quarterly Journal of Economics* 90:591–597.

Strauss, J. 1986. "The Theory and Comparative Statics of Agricultural Household Models: A General Approach." In I. Singh, L. Squire, and J. Strauss, eds. *Agricultural Household Models: Extensions, Applications, and Policy*. Baltimore: Johns Hopkins University Press, pp. 71–91.

Thijssen, G.J. 1988. "Estimating a Labour Supply Function of Farm Households." *European Review of Agricultural Economics* 15:67–78.

Vakis, R., E. Sadoulet, and A. de Janvry. 2003. "Measuring Transactions Costs from Observed Behavior: Market Choices in Peru." http://are.berkeley.edu/~sadoulet/papers/TC-wp.pdf.

Vance, C., and J. Geoghegan. 2004. "Modeling the Determinants of Semi-Subsistent and Commercial Land Uses in an Agricultural Frontier of Southern Mexico: A Switching Regression Approach." *International Regional Science Review* 27:326–347.

Venables, W.N., and B.D. Ripley. 2002. *Modern Applied Statistics with S*, 4th ed. New York: Springer.

Yee, T.W., and C.J. Wild. 1996. "Vector Generalized Additive Models." *Journal of the Royal Statistical Society, Series B, Methodological* 58:481–493.

Zawadzki, W. 1994. "Infrastruktura Techniczna - Glowny Warunek Wielofunkcyjnego Rozwoju Obszarow Wiejskich." Tech. Rep. 365, IERiGZ, Warsaw.

Kapitel 8

Taxation of the Farm Household and Imperfect Labor Markets: Theoretical Results and Empirical Evidence from Polish Farm-Household Data

Christian H.C.A. Henning und Arne Henningsen

zur Veröffentlichung eingereicht beim: Journal of Development Economics

Abstract

Recent analyses of agricultural tax policies show how standard presumptions in public finance literature change when market imperfections are taken into account (Hoff et al., 1993). This paper provides a theoretical analysis as well as an econometric estimation of an interdependent farm household model (FHM) approach considering imperfect labor markets and various tax policy instruments. Comparative static analysis supports the results of recent studies showing that neither standard nor land taxes are theoretically superior when compared to agricultural tax instruments. However, empirical estimation with individual household data from Poland partly confirms standard presumptions, i.e. even when markets are imperfect income and consumption taxes imply negligible production adjustments.

JEL classification: H21, Q12, J22, D43, C51

Key words: agricultural taxation, farm household model, labor market, market imperfection, Poland

8.1 Introduction

Since Ramsey's (1927) classical paper was published, a central result in public finance theory is that optimal taxation policies correspond to a combination of consumption (value-added) and income taxes, assuming that those taxes imply no production effects. However, in many countries, taxation of farm households differs from the standard tax scheme applied to the non-peasant economy (Newbery and Stern, 1987; Ahmad and Stern, 1991). Differences often include both different tax levels and different instruments of taxation for farm households. There is extensive literature on the analysis of so-called agricultural tax tools applied to peasant households as surrogates for the standard income or profit taxes (Rao, 1989; Burgess and Stern, 1993).

Explanations of different tax schemes for the peasant and non-peasant economy are manifold. A prominent justification of the application of specific agricultural taxes, especially in developing and transition countries, refers to specific structural characteristics of small family farms like internal transfers of self-produced consumption goods or lack of appropriate accounting systems. These characteristics limit the information available to tax authorities, which in turn restrains tax instruments that can be used at reasonable administrative costs (Hoff et al., 1993; Besley, 1993). Moreover, specific agricultural tax systems are often explained by lobbying activities. For instance, Hirschman (1964) explains that a land tax is often not applied despite its economic efficiency with the political power of land owners.

While a final explanation of specific agricultural tax systems remains unclear, it is a standard presumption in public finance literature that when administrative costs are neglected, standard taxes and a land tax are economically efficient, while so-called agricultural taxes like market surplus, output, or input taxes induce production distortions and, hence, are economically inefficient. Diamond and Mirrless (1971) point out that production efficiency is desirable within an optimal taxation system, although a full Pareto optimum is not achieved, because commodity taxes imply that marginal rates of substitution are not equal to the marginal rates of transformation. Accordingly, it is a standard conclusion that in contrast to market surplus, output, and input taxes, a land tax is efficient, implying no production distortions (Newbery, 1987).

Hence, given the trade-off between political and administrative feasibility and economic efficiency, optimal agricultural tax design is intensively discussed at both policy-making and academic levels (Rao, 1989; Burgess and Stern, 1993; Newbery, 1987; Atkinson, 1987; Heady and Mitra, 1987; Sah and Stiglitz, 1987; Newbery and Stern, 1987; Ahmad and Stern, 1991; Munk, 1980).

This discussion has been further augmented by recent studies of agricultural tax policies incorporating imperfect market institutions. In particular, these studies challenge standard presumptions in public finance literature regarding economic efficiency of standard and land taxes when compared to agricultural taxes. For example, assuming missing markets for risk Hoff and Stiglitz (1993) demonstrate that an optimal agricultural tax policy mix includes an output tax, contradicting standard presumptions. Taking missing or imperfect markets into consideration results in two major implications for tax policy analysis. First, market failure generally implies second-best policies. Thus, general efficiency theorems derived under a first-best scenario, i.e. assuming perfect markets, no longer apply. For example, assuming that small farms have limited access to credit markets and therefore can only buy a limited amount of fertilizer implies that input subsidisation can be Pareto-efficient even if environmental pollution is involved (Hoff, 1993). Second, the economic environment, i.e. the response of the farm household to different tax policies, changes when missing or imperfect markets are taken into account.

Therefore, since preexisting studies on agricultural taxes generally assume perfect markets (e.g. Newbery, 1987; Atkinson, 1987; Heady and Mitra, 1987; Sah and Stiglitz, 1987; Newbery, 1987; Ahmad and Stern, 1991), the new approach suggested by Hoff et al. (1993) certainly makes an important contribution to the agricultural tax literature. However, these new studies are merely theoretical and focus mainly on the first point. Moreover, they only analyse land and other specific agricultural taxes, while standard tax policies are mainly neglected.[1] Finally, these studies assume totally missing markets while the more realistic case of existing but imperfect markets has not yet been considered. However, responses of farm households to price and income changes significantly differ between totally missing and existing but imperfect markets (Henning and Henningsen, 2005b,a).

Thus, overall it is still fair to say that a comprehensive empirical analysis of taxation of farm households under imperfect and missing markets does not yet exist in the literature. This is regrettable, because beyond a theoretical analysis, an empirical analysis would be desirable to address the question to what extent or under what conditions the new theoretically identified counterintuitive effects of market imperfections translate into significant quantitative effects. The latter is especially relevant for practical design of peasant tax schemes (Chambers and Lopez, 1987, p. 373).

In this context the paper provides a theoretical and empirical analysis of tax policies for farm households facing imperfect labor markets. The theoretical model corresponds to an interdependent farm household model (FHM) already found in the literature (Sadoulet et al., 1998; Henning, 1994; Sonoda and Maruyama, 1999). However, the applied approach differs regarding the modeling of imperfect labor markets.

We use an FHM approach including fixed (FTC), proportional (PTC), and variable marginal (VTC) transaction costs as well as observed heterogeneity of labor by assuming a non-linear labor income function for off-farm labor supply and a non-linear labor cost function for hired on-farm labor. Technically, this approach corresponds to a model that has been derived to estimate price responses of farm households in imperfect labor markets (Henning and Henningsen, 2005b,a). Based on Henning and Henningsen (2005a), major shortcomings inherent in earlier work on taxation of the farm household are overcome in this paper[2].

Overall, our modeling strategy has the following advantages: (i) The approach allows a consistent theoretical and empirical analysis of standard and agricultural tax policies for perfect, imperfect, and missing labor markets. (ii) It provides a theoretical and empirical comparative static analysis of various tax instruments, i.e. an income, a value-added, a market

[1]Standard tax policies are analyzed by Chambers and Lopez (1987), presenting a very sophisticated analysis of the implications of income, profit, and consumption taxes on the economic decisions of financially constrained farm households within a dynamic farm household model (FHM). However, agricultural tax policies are not analyzed in their paper.

[2]The basic concept of the applied FHM approach including imperfect labor markets and tax policies has originally been developed by Henning (1995) and empirically applied to Polish Farm Household data by Henning (1998). Further detailed theoretical and empirical analyses of agricultural tax policies in Poland have been undertaken (Glauben, 2000; Glauben et al., 2003), based on the earlier work of Henning (1995, 1998). However, all of this earlier work suffers from some shortcomings in the theoretical concept and econometric estimation of labor supply and demand functions. These shortcomings finally have been overcome in Henning and Henningsen (2005a).

surplus, an input and a land tax, as well as a wage tax for off-farm employment. (iii) Unlike most empirical analyses of interdependent FHMs using reduced forms of a non-separable FHM, we consistently estimate a full non-separable FHM based on flexible functional forms on the production and consumption side. (iv) Compared to approaches considering only FTC and PTC (Key et al., 2000; Sadoulet et al., 1998), non-separability occurs in our approach not only if households are autarkic but also when households actually participate in labor markets. (v) We provide a simple empirical test of market imperfection.

8.2 Theoretical Model

To concentrate on the role of tax policies and labor market constraints, we construct a static model that ignores some aspects of farmers' decisions, notably (price) risk (Finkelshtain and Chalfant, 1991; Fafchamps, 1992) and credit constraints (Chambers and Lopez, 1987). The model covers perfect, imperfect, and totally missing labor markets. The farm household is assumed to maximise utility subject to a technology (8.2), time (8.3), and a 'tax-corrected' budget (8.4) constraint. Therefore, farm households solve the following maximization problem:

$$\max_{x,c} U(c) \tag{8.1}$$

subject to

$$G(x, r) = 0 \tag{8.2}$$

$$T_L - |X_L| + X_L^h - X_L^s - C_L \geq 0 \tag{8.3}$$

$$\begin{aligned} (1+\tau_{vat})P_m C_m + P_a C_a \quad \leq \quad & (1-\tau_y)\Big\{(1-\tau_{ms})\Big[P_c X_c + P_a(X_a - C_a)\Big] + P_a C_a \\ & -(1+\tau_v)P_v|X_v| - g(X_L^h) + (1-\tau_w)f(X_L^s) + E\Big\} - \tau_g V_g \end{aligned} \tag{8.4}$$

Here $U(c)$ is the farm household's utility function, which is assumed to be monotonically increasing and strictly concave. c is a vector of consumption goods consisting of market commodities (C_m), self-produced agricultural goods (C_a), and leisure (C_L). Production technology is represented by a multi-output, multi-input production function (8.2), which is preconditioned to be well behaved in the usual sense (Lau, 1978a). Here x is a vector of production goods, expressed as netputs, and r is a vector of quasi-fixed factors. The farm household is presumed to produce market ($X_c > 0$) and home-consumed ($X_a > 0$) agricultural goods using variable inputs ($X_v < 0$), labor ($X_L < 0$), and the quasi-fixed factors land (R_g) and capital (R_k).

The farm household faces a time constraint (8.3), where T_L denotes the total time available. $|X_L| = X_L^f + X_L^h$ is the total of on-farm labor time subdivided into family labor (X_L^f) and hired labor (X_L^h), and X_L^s denotes off-farm family labor. In general, there are four possible regimes of labor market participation. First, the household simultaneously sells family labor and hires non-family labor. Second, farmers neither sell nor hire labor (autarky). Third and fourth, they either sell or hire labor.

The budget constraint (8.4) states that a household's ('tax-corrected') consumption expenditures (left-hand side) must not exceed its ('tax-corrected') monetary income (right-hand side). The household may receive income from farming and from off-farm employment. In addition, it receives ($E > 0$) or pays ($E < 0$) transfers, which are determined exogenously. Here, P_i, $i \in \{m, a, c, v\}$ denotes an exogenous consumer and producer price before tax, and τ_j are the parameters of tax policies to be analyzed. Legally, total monetary expenditures are subject to a value-added tax τ_{vat}. However, tax authorities cannot observe internal transfer of self-produced agricultural goods for farm households. Thus, only the expenditures for market commodities ($P_m C_m$) are subject to the value-added tax. The base of the income tax (τ_y) is the household's monetary income, including off-farm labor income ($f(X_L^s)$), transfers (E), and profits from farming ($P_c X_c + P_a X_a - P_v|X_v| - g(X_L^h)$), where $g(X_L^h)$ denotes hired labor costs (see below). In the absence of record keeping, farm income is often not taxable and, therefore, only incomes from off-farm employment can be taxed by a wage tax (τ_w). Similarly, market surplus, input, or land taxes are applied as surrogates for an income tax. The base of the market surplus tax (τ_{ms}) is revenue from sales of agricultural goods ($P_c X_c + P_a(X_a - C_a)$). Expenditures for commercial inputs ($P_v|X_v|$) such as fertilizers and chemicals are subject to the input tax (τ_v) and the market value of land (V_g) is subject to a land tax (τ_g).

A special emphasis is given to the modeling of labor markets. It is well recognized in the literature that rural labor markets are often plagued by transaction costs. These are considered as fixed (FTC) and proportional (PTC) transaction costs in existing studies (Key et al., 2000; Vakis et al., 2003). Although the concept of FTCs and PTCs appears appealing at first glance, Henning and Henningsen (2005b,a) show that additionally variable marginal (= non-proportional variable) transaction costs (VTC) are generally conceivable and should not be excluded ex ante. To formally include VTCs as well as FTCs and PTCs in our model, we denote total variable transaction costs (VTC + PTC) of off-farm employment by $TC_v^s(X_L^s, z_v^s)$. In general, transaction costs are not observable (Key et al., 2000); however, some factors that explain these transaction costs can be observed, where z_v^s indicates the factors explaining variable transaction costs of the farm household for selling labor.

Furthermore, it is reasonable to assume that family members have heterogeneous skills to work off-farm. This (observed) heterogeneity implicates different wage rates. If we further assume that family labor is homogeneous regarding farm work, profit maximization implies that the order in which family members work off-farm corresponds to their skill levels. Hence, the marginal off-farm wage is a step-wise decreasing function of off-farm labor supply. To facilitate theoretical and empirical analyses we approximate the step-wise labor wage function by a continuous function. Subtracting marginal transaction costs we get following effective marginal labor wage function:

$$P_L^s(\overline{P}_L, X_L^s, z_L^s, z_v^s) = \overline{P}_L + b^s(X_L^s, z_L^s) - \frac{\partial TC_v^s(X_L^s, z_v^s)}{\partial X_L^s}, \tag{8.5}$$

where $\overline{P}_L$ denotes the average regional labor wage, z_L^s denotes the factors explaining heterogeneity of the quality of family labor regarding off-farm work, and $b^s(X_L^s, z_L^s)$ denotes the up or down shift of the average labor wage observed by the farm household. According to our above expositions we expect that b^s is non-increasing in labor supply.

Taking observable heterogeneity and transaction costs into account, the effective revenues from off-farm employment are a function of supplied labor time:

$$f(\overline{P}_L, X_L^s, z_L^s, z_v^s, z_f^s) = \overline{P}_L X_L^s + \int_0^{X_L^s} b(X_L^s, z_L^s)\, \mathrm{d}\, X_L^s - TC_v^s(X_L^s, z_v^s) - Y^s TC_f^s(z_f^s), \tag{8.6}$$

where Y^s equals one, if $X_L^s > 0$ and zero otherwise, $TC_f^s(z_f^s)$ denotes fixed transaction costs, and z_f^s are factors explaining fixed transaction costs of supplying off-farm labor.

Assuming variable marginal transaction costs and observed heterogeneity also for hiring labor, the effective marginal wage rate for on-farm labor and the effective labor cost function can be derived analogously:

$$P_L^h(\overline{P}_L, X_L^h, z_L^h, z_v^h) = \overline{P}_L + b^h(X_L^h, z_L^h) + \frac{\partial TC_v^h(X_L^h, z_v^h)}{\partial X_L^h}, \tag{8.7}$$

$$g(\overline{P}_L, X_L^h, z_L^h, z_v^h, z_f^h) = \overline{P}_L X_L^h + \int_0^{X_L^h} b^h(X_L^h, z_L^h)\, \mathrm{d}\, X_L^h + TC_v^h(X_L^h, z_v^h) + Y^h TC_f^h(z_f^h), \tag{8.8}$$

where $b^h(X_L^h, z_L^h)$ indicates the up or down shift of the average regional labor wage observed by the farm household, $TC_v^h(z_v^h)$ and $TC_f^h(z_f^h)$ denote total variable and fixed transaction costs for hiring on-farm labor, respectively, z_v^h and z_f^h are factors explaining total variable and fixed transaction costs, respectively, and Y^h equals one, if $X_L^h > 0$ and zero otherwise.

As mentioned above, our theoretical framework is applicable for several kinds of labor market imperfections including FTCs, PTCs, and VTCs, as well as observable heterogeneity. It follows directly from the definition of $f(.)$ and $g(.)$ above that curvature properties of these functions correspond to market imperfection. For example, in the absence of heterogeneity and VTCs, both functions become linear:

$$f(.) = \Big(\overline{P}_L - TC_p^s + b_p^s(z_L^s)\Big) X_L^s - Y^s TC_f^s(z_f^s) \tag{8.9}$$

$$g(.) = \Big(\overline{P}_L + TC_p^h + b_p^h(z_L^h)\Big) X_L^h + Y^h TC_f^h(z_f^h) \tag{8.10}$$

where TC_p^s and TC_p^h are proportional transaction costs per unit of labor, and $b_p^s(z_L^s)$ and $b_p^h(z_L^h)$ denote the up or down shift of the average regional labor wage observed by the farm household for supplying and hiring labor, respectively. Hence, once households participate in labor markets, marginal off-farm income or marginal costs for hired labor are equal to the exogenously given regional wage rate ($\overline{P}_L$), corrected for proportional transaction costs (TC_p^s and TC_p^h) as well as for individual wage shifters (b_p^s and b_p^h). Thus, if households participate in one of the labor markets the farm household model becomes separable and delivers standard microeconomic comparative static results (Sadoulet et al., 1998). Of course, if fixed or proportional transaction costs are too high, households may still abstain from the labor market and stay autarkic, implying a non-separable FHM (Key et al., 2000). In contrast, when labor markets are assumed to be imperfectly competitive due to heterogeneity or VTCs, both functions become non-linear. In this case, the shadow price of family labor (P_L^*) is endogenously determined and production and consumption decisions are simultaneously determined by the solution of the utility maximization problem (8.1) to (8.4). Note that in contrast to FTCs or PTCs, heterogeneity or VTCs result in non-separability although households actually participate in labor markets.

As FTCs create discontinuities in the $f(.)$ and $g(.)$ functions, solutions to the maximization problem (8.1) to (8.4) cannot be found by simply solving the first-order conditions. Thus, we follow Key et al. (2000) and decompose the solution into two steps. First, we solve the maximization problem for each of the four labor market regimes (Y^h and Y^s), and, second, we choose the regime that leads to the highest level of utility. Thus, assuming there is an interior solution for a given labor market regime (Y^h and Y^s), the optimal quantities of consumption and production goods and the allocation of time are determined ($\lambda, \phi, \mu > 0$; $C_m, C_a, C_L, X_c, X_a > 0$; $X_L, X_v < 0$; $X_L^s > 0$ if $Y^s = 1$ and $X_L^s = 0$ otherwise, and $X_L^h > 0$ if

$Y^h = 1$ and $X_L^h = 0$ otherwise).

$$\frac{\partial U(.)}{\partial C_i} - \lambda P_i^{c*} = 0 \quad i \in \{m, a, L\} \tag{8.11}$$

$$\phi \frac{\partial G(.)}{\partial X_i} + \lambda P_i^{p*} = 0 \quad i \in \{c, a, v, L\} \tag{8.12}$$

$$\frac{\partial f^*(.)}{\partial X_L^s} = P_L^* \quad \text{if} \quad Y^s = 1 \tag{8.13}$$

$$\frac{\partial g^*(.)}{\partial X_L^h} = P_L^* \quad \text{if} \quad Y^h = 1 \tag{8.14}$$

$$\sum_{i \in \{c,a,v\}} P_i^{p*} X_i - g^*(X_L^h) + f^*(X_L^s) + E^* - V_g^* - \sum_{i \in \{m,a\}} P_i^{c*} C_i = 0 \tag{8.15}$$

$$G(x, r) = 0 \tag{8.16}$$

$$T_L + X_L + Y^h X_L^h - Y^s X_L^s - C_L = 0 \tag{8.17}$$

Here $\lambda, \phi > 0$ are Lagrangian multipliers associated with the budget and the technology constraints, respectively. For the non-separable FHM we denote $P_L^* = \mu/\lambda$ the unobservable internal wage, where μ is the Lagrangian multiplier associated with the time constraint. Moreover, we introduce decision prices, P_j^{c*} and P_j^{p*}, for consumer and producer goods, respectively. In particular, $P_m^{c*} = (1 + \tau_{vat})P_m$ and $P_a^{c*} = (1 - (1 - \tau_y)\tau_{ms})P_a$ are the tax-corrected decision prices of consumption goods.[3] Analogously, we define decision prices of production goods by $P_j^{p*} = (1 - \tau_y)(1 - \tau_{ms})P_j$; $j \in \{a, c\}$, and $P_v^{p*} = (1 - \tau_y)(1 + \tau_v)P_v$. Decision prices of labor and leisure differ between the separable and non-separable FHM. For the non-separable model these prices are equal to the internal wage rate ($P_L^{p*} = P_L^{c*} = P_L^*$). As for the separable model farm households either supply or demand labor, the decision prices of labor and leisure are exogenously given by the tax-corrected off-farm wage in the first case ($P_L^{p*} = P_L^{c*} = (1 - \tau_y)(1 - \tau_w)P_L^s$) and the tax-corrected wage of hired labor in the second case $\left(P_L^{p*} = P_L^{c*} = (1 - \tau_y)P_L^h\right)$.

In addition, we use the following definitions in (8.11) - (8.17): $f^*(.) = (1 - \tau_y)(1 - \tau_w)f(.)$, $g^*(.) = (1 - \tau_y)g(.)$, $E^* = (1 - \tau_y)E$ and $V_g^* = \tau_g V_g$.

[3] Note that the decision price for C_a depends on the net-trade position of the farm-household, where the above defined decision prices result under the assumption that the household is a net-supplier ($X_a > C_a$).

8.3 Comparative Static

In general, tax policies have an impact on farm households' allocation decisions, because tax policies determine households' decision prices and income. Thus, comparative static of tax policies basically corresponds to the sum of comparative static effects of induced price and income changes. However, the effect of various taxes on decision prices and total household income and the household's response to these changes depends on the institutional environment and the labor market regime in which the farm household operates. Given the well-known result of the FHM literature that price responses of the farm household differ between perfect and imperfect markets (de Janvry et al., 1991, or more recently Henning and Henningsen, 2005b,a), one would assume that this also generally holds for the responses to tax policies. To formally prove this intuition, we derive the comparative static of various tax policies from the equation system (8.11) to (8.17) for the three different institutional settings. To simplify the analysis we assume a specific corresponding labor market regime for each institutional setting.

For perfect labor markets we assume that the household supplies labor but does not demand labor ($Y^s = 1, Y^h = 0$). Under this assumption, the decision price of labor is exogenously determined by (8.9) and (8.13), while the off-farm labor supply is residually determined by (8.17). Thus, for the separable FHM the comparative static of various tax policies simply corresponds to the sum of standard comparative static effects of induced changes of exogenous decision prices and total income.[4]

In contrast, assuming imperfect labor markets, the decision price of labor is endogenously determined by (8.11) - (8.17). Again, the comparative static of tax policies depends on the actual labor market regime. In this case we assume that the farm household simultaneously supplies off-farm labor and demands on-farm labor ($Y^s = Y^h = 1$). Furthermore, the autarky regime ($Y^s = Y^h = 0$) corresponds to the case of totally missing labor markets. Note that in contrast to Sadoulet et al. (1998), both regimes are prominent in our farm survey of Mid-West Poland. In contrast to perfect labor markets, both imperfect and missing markets imply an interdependent FHM. Following the standard FHM literature (de Janvry et al., 1991) comparative static of an interdependent FHM can be decomposed into the following two components:

$$\frac{\mathrm{d}\,Q}{\mathrm{d}\,\tau_j} = \left.\frac{\partial Q}{\partial \tau_j}\right|_{P_L^*=\text{const.}} + \frac{\partial Q}{\partial P_L^*}\frac{dP_L^*}{d\tau_j} \tag{8.18}$$

[4]However, on the consumption side an additional income term has to be taken into account when compared to standard microeconomic models (Strauss, 1986; Henning, 1994).

The first term (direct component) on the right-hand side represents the supply or demand reactions to changes in the exogenous decision prices induced by a tax policy, assuming a constant endogenous labor price (P_L^*). The second term (indirect component) represents the adjustments to the changes in the internal wage rate caused by the same tax policy.

To determine the indirect component of the non-separable version, we derive the tax-induced shadow price adjustment applying the implicit function theorem to the time constraint (8.3) (de Janvry et al., 1991):

$$\frac{\mathrm{d}\,P_L^*}{\mathrm{d}\,\tau_j} = -\frac{\dfrac{\partial X_L}{\partial \tau_j} + \dfrac{\partial X_L^h}{\partial \tau_j} - \dfrac{\partial X_L^s}{\partial \tau_j} - \dfrac{\partial C_L}{\partial \tau_j}}{\dfrac{\partial X_L}{\partial P_L^*} + \dfrac{\partial X_L^h}{\partial P_L^*} - \dfrac{\partial X_L^s}{\partial P_L^*} - \dfrac{\partial C_L^H}{\partial P_L^*}} = -\frac{\dfrac{\partial X_L}{\partial \tau_j} + \dfrac{\partial X_L^h}{\partial \tau_j} - \dfrac{\partial X_L^s}{\partial \tau_j} - \dfrac{\partial C_L}{\partial \tau_j}}{\dfrac{\partial^2 \Pi}{\partial P_L^{*2}} + \left(\dfrac{\partial^2 g}{\partial {X_L^h}^2}\right)^{-1} - \left(\dfrac{\partial^2 f}{\partial {X_L^s}^2}\right)^{-1} - \dfrac{\partial^2 e}{\partial P_L^{*2}}} \tag{8.19}$$

The numerator on the right-hand side represents the change in the time allocation due to an increasing tax rate. Here, $\partial X_L/\,\partial \tau_j = \sum_{i\in\{c,a,v\}} (\partial X_L/\,\partial P_i^*)\,(\partial P_i^*/\,\partial \tau_j)$ denotes tax-induced on-farm labor adjustment, $\partial C_L/\,\partial \tau_j = \sum_{i\in\{m,a\}} (\partial C_L/\,\partial P_i^*)\,(\partial P_i^*/\,\partial \tau_j)$ denotes the tax-induced adjustment of Marshallian leisure demand and $\partial C_L^H/\partial P_L^*$ denotes the Hicksian leisure demand response to changed labor prices. Moreover, $\partial X_L^h/\partial \tau_j$ and $\partial X_L^s/\partial \tau_j$ are the direct labor market reactions to changed tax policies.[5]

To facilitate the comparative static analysis we follow the standard procedure in FHM literature (de Janvry et al., 1991) and derive the comparative static from a dual restricted profit function $\Pi(p_p^*, r)$ and a dual expenditure function $e(p_c^*, U^0)$, where p_p^* and p_c^* are the tax-corrected price vectors of the production and consumption goods, respectively. Given the convexity of $\Pi(.)$ and the concavity of $e(.)$ in prices, the denominator is always positive as long as it is assumed that $g(.)$ is convex in X_L^h and $f(.)$ is concave in X_L^s. Substituting equation (8.19) into expression (8.18) yields the comparative statics of the FHM.

Based on equations (8.18) and (8.19) we derive the complete comparative static for all tax instruments mentioned above.[6] The results are summarized in table 8.1. In this table we present only comparative statics of tax policies assuming perfect and imperfect markets, because in qualitative terms, though not in quantitative terms, results for missing markets

[5] Direct labor market reactions only result for an income and a wage tax, since only these taxes directly affect the tax-corrected labor income and cost functions, $f(.)$ and $g(.)$, respectively (Glauben, 2000). In particular, it holds: $\left.\frac{\partial X_L^s}{\partial \tau_j}\right|_{P_L^*=\text{const.}} = -\frac{P_L^*}{(1-\tau_j)} \Big/ \frac{\partial^2 f^*(.)}{\partial {X_L^s}^2} < 0$; $(\tau_j | j = y, w)$, and $\left.\frac{\partial X_L^h}{\partial \tau_y}\right|_{P_L^*=\text{const.}} = -\frac{P_L^*}{(1-\tau_y)} \Big/ \frac{\partial^2 g^*(.)}{\partial {X_L^h}^2} > 0$.

[6] All designed tax policies are considered as *alternative* tax instruments. Hence, it is assumed that the respective tax is the only tax policy applied to the farm household. A detailed documentation of the comparative static is available from the authors upon request.

generally correspond to imperfect markets. One exception is the wage tax, because a wage tax has no impact on the farm household's equilibrium in the case of missing labor markets.

Table 8.1: Theoretical effects of taxes

	Non-separable model						Separable model					
	τ_y	τ_w	τ_{ms}	τ_v	τ_g	τ_{vat}	τ_y	τ_w	τ_{ms}	τ_v	τ_g	τ_{vat}
X_c	?	+	?	?	+	?	0	+	-	-	0	0
X_a	?	+	?	?	+	?	0	+	-	-	0	0
$\|X_L\|$	?	+	?	?	+	?	0	+	-	(-)	0	0
$\|X_v\|$	?	(+)	?	?	(+)	?	0	(+)	-	-	0	0
C_m	(-)	(-)	?	?	(-)	?	(-)	(-)	(-)	-	-	-
C_a	(-)	(-)	?	?	(-)	?	(-)	(-)	?	-	-	?
C_L	?	?	?	?	?	?	?	?	(-)	-	-	?
X_L^{sn}	?	?	?	?	+	?	?	?	(+)	(+)	+	?
X_L^s	?	?	?	?	+	?						
X_L^h	?	-	?	?	-	?						
P_L^*	-	-	?	?	-	?						

Notes: It is assumed that goods are not inferior, technologies are not regressive, and farmers are net suppliers of labor.

0	=	clear, no effect;
+/-	=	clear, increase/decrease;
(+)/(-)	=	unclear, but most likely an increase/decrease (assuming labor and variable inputs are complements, and consumption goods are net-substitutes);
?	=	unclear.

The results presented in table 8.1 support major implications of recent agricultural tax studies (Hoff et al., 1993) highlighting the importance of the institutional environment for a farm household's responses to tax policies. In particular, when markets are missing or imperfect, standard taxes, like the income or consumption tax, imply adjustments on the production side, contradicting classical presumptions in public finance. Analogously, contrary to standard presumptions, a land tax implies production adjustments for imperfect or missing labor markets. Moreover, while assuming perfect markets, market surplus and input taxes have a clear negative impact on farm production and consumption. However, in the case of missing or imperfect markets these impacts are no more theoretically determined. Interestingly, while Hoff et al. (1993) derive these non-standard results for imperfect risk and credit markets, we get corresponding results for imperfect labor markets. Additionally, we could generalize these non-standard results to a more common case of market failure, i.e.

households actually do participate in markets, but market participation is plagued by transaction costs or heterogeneity. Of course, in real rural economies often all of these markets are missing or at least partly imperfect. Therefore, it will be interesting to analyse how simultaneous imperfections in all of these markets will affect farm households' responses to tax or price policies. We consider our suggested FHM approach as a promising methodological framework for this analysis. However, we leave this interesting topic for future research.

How can these partly counterintuitive results be explained? For missing markets, these follow directly from the well-known logic of the internal equilibrium of interdependent FHM (de Janvry et al., 1991). Any change of exogenous parameters, prices, income or other variables induces an additional change of the internal shadow price equilibrium. Furthermore, these shadow price changes imply additional production and consumption adjustment that do not occur when markets are perfect, i.e. all prices are exogenous. Given the expositions above, it follows quite plainly that these specific internal shadow price adjustments also occur when markets do exist and households actually do participate in these markets, but markets are imperfect, i.e. households observe variable marginal transactions or heterogeneity (see also Henning and Henningsen, 2005b,a).

Formally, for most tax policies the difference between comparative statics of the separable and the non-separable FHM can be seen directly from (8.18), because the direct component of the non-separable FHM generally equals the corresponding comparative static of the separable model. Exceptions are the labor wage and income tax, because these taxes directly change the exogenous decision price of labor in the separable model and, therefore, the comparative static of the separable model is not equal to the direct component of the non-separable model.

For example, assuming perfect markets, an income tax simply changes the nominal decision prices of all production goods proportionally, leaving relative decision prices unchanged. Thus, given the homogeneity of the profit function, no impact on farm production is implied. In contrast, an income tax implies a direct change of all exogenous production decision prices for missing or imperfect markets, while the endogenous labor wage is adapted according to specific internal adjustment processes. Hence, an income tax generally implies a change in the labor price in relation to other exogenous production decision prices, inducing corresponding production adjustments. Note further that the direct component of comparative static effects in (8.18) only captures production adjustments due to implied changes of exogenous decision prices, while adjustment processes due to the implied endogenous labor wage change are captured by the indirect component. Hence, it follows directly that for an income tax the comparative static effect of the separable model does not equal the direct component of the comparative static effect of the interdependent model.

Note that the impact of the wage tax on the farm household's production and consumption crucially depends on the labor market regime, i.e. the household's participation in off-farm labor market. Despite specific institutional framework conditions the wage tax has an impact on farm production and consumption only when farms participate in off-farm labor markets. Therefore, as far as the wage tax is concerned, results in table 8.1 are mainly determined by assumed labor market regimes and much less a characteristic implication of specific labor market conditions. Note further that off-farm labor participation is significantly determined by relative land-family-labor endowment, where typically small family farms are net-suppliers of off-farm labor. Hence, the impact of the wage tax should also be systematically related to this specific family type.

Overall, our theoretical results underline that the organisation of rural labor markets matters for the positive analysis of taxation of farm households. In particular, responses to tax policies significantly differ between perfect and imperfect or missing markets. This implies that the central standard presumption of public finance literature has to be revised. So far our comparative static analyses focus on positive analyses of a farm household's responses to various tax policies, while we have not yet discussed implications for an optimal taxation system from a normative welfare economic perspective. Nevertheless, given the fact that the economic environment, i.e. farm responses to tax policies, is central for deriving any optimal taxation system, two general conclusions regarding optimal taxation of the farm household can be drawn from our theoretical comparative static analysis. First, while a general normative priority of tax policies can be defined on the ground of theoretical comparative static results for perfect markets, this is no longer possible if missing or imperfect markets are considered because all tax policies induce allocation responses on the consumption and production side. Therefore, when markets are partially imperfect both theoretical optimal taxation analysis as well as practical design of optimal agricultural taxation systems require a quantitative assessment of farm household's responses to various tax policies. Second, for imperfect markets optimal taxation policy analyses correspond to second-best policy analyses, where existing imperfect market institutions have to be taken into account. Hence, optimal tax policies depend on the specific institutional environment. As far as imperfect markets are concerned, beyond tax policy reform institutional reform is a further important political option to be considered. Thus, given the fact that the achievement of perfect markets is often an unrealistic goal for many transition and developing countries, the modeling of gradual institutional reforms is another interesting topic for future research.

However, given the general importance of quantitative empirical analyses of farm responses to various tax policies, and given the fact that empirical estimations of these responses have rarely been undertaken in the literature, this paper is devoted to this important endeavor. We provide an econometric estimation of our FHM approach based on farm household data from Mid-West Poland. According to our theoretical analysis the empirical analysis focuses on the quantitative assessment of tax responses under different labor market conditions and on an empirical test of market imperfection.

8.4 Empirical Specification

The farm household model is specified as follows: Production decisions are represented by a multi-output, multi-input profit function from the symmetric normalized quadratic (SNQ)[7] form (Diewert and Wales, 1987, 1992; Kohli, 1993). We apply the method proposed by Koebel et al. (2000, 2003) to ensure global convexity (see below). The consumption decisions of the farm households are specified by an 'Almost Ideal Demand System' (AIDS) (Deaton and Muellbauer, 1980). To allow labor market imperfection in terms of FTCs, PTCs, and VTCs, as well as (observed) heterogeneity, we assume a quadratic form for the labor cost (g) and labor income (f) function.[8]

The econometric estimation of the proposed model is carried out in four steps. First, we determine the internal wage of the household by estimating the shadow price of labor on the farm. We follow Lopez (1984), who estimated a restricted profit function with labor as a quasi-fixed factor. Assuming constant returns to labor Lopez (1984) could directly derive shadow prices of labor from the estimated profit function. In this case the SNQ profit function is defined as follows:

$$\begin{aligned}\Pi(p_n^{p*}, r_n, X_{Ln}) &= X_{Ln}\Pi^L(p_n^{p*}, r_n) \\ &= X_{Ln}\left(\begin{array}{l}\displaystyle\sum_{i\in\{c,a,v\}}\alpha_i P_{in}^{p*} + \frac{1}{2}w_n^{-1}\sum_{i\in\{c,a,v\}}\sum_{j\in\{c,a,v\}}\beta_{ij}P_{in}^{p*}P_{jn}^{p*} \\ \displaystyle+\sum_{i\in\{c,a,v\}}\sum_{j\in\{g,k\}}\delta_{ij}P_{in}^{p*}R_{jn} + \frac{1}{2}w_n\sum_{i\in\{g,k\}}\sum_{j\in\{g,k\}}\gamma_{ij}R_{in}R_{jn}\end{array}\right)\end{aligned} \tag{8.20}$$

[7] This functional form is also traded under the name of 'symmetric generalized McFadden function' (Diewert and Wales, 1992).

[8] The quadratic form can be interpreted as a second-order approximation of the true labor cost and income functions.

and the corresponding netput equations are

$$X_{in}(p_n^{p*}, r_n, X_{Ln}) = X_{Ln} \begin{pmatrix} \alpha_i + w_n^{-1} \sum\limits_{j\in\{c,a,v\}} \beta_{ij} P_{jn}^{p*} \\ -\frac{1}{2}\theta_i w_n^{-2} \sum\limits_{j\in\{c,a,v\}} \sum\limits_{k\in\{c,a,v\}} \beta_{jk} P_{jn}^{p*} P_{kn}^{p*} \\ + \sum\limits_{j\in\{g,k\}} \delta_{ij} R_{jn} + \frac{1}{2}\theta_i \sum\limits_{j\in\{g,k\}} \sum\limits_{k\in\{g,k\}} \gamma_{jk} R_{jn} R_{kn} \end{pmatrix} \tag{8.21}$$

where n indicates the observation (household), Π is the profit function, Π^L is the profit function per unit of labor, X_{Ln} is the labor deployed on the farm and $w_n = \sum_{i\in\{c,a,v\}} \theta_i P_{in}^{p*}$ is a factor to normalize prices, where $\theta_i = \sum_n P_{in}^{p*} |X_{in}| / \sum_n \sum_{j\in\{c,a,v\}} P_{jn}^{p*} |X_{jn}|$; $i \in \{c, a, v\}$ are the weights of the individual netput prices. Further, $p_n^{p*} = (P_{an}^{p*}, P_{cn}^{p*}, P_{vn}^{p*})$ indicates the netput prices and X_{in}; $i \in \{c, a, v\}$ denotes the netput quantities. $r_n = (R_{gn}, R_{kn})$ represents the quasi-fixed factors land (R_g) and capital (R_k), and α, β, δ, γ are the parameters to be estimated. To identify all β coefficients, we impose the restrictions $\sum_{j\in\{c,a,v\}} \beta_{ij} \overline{P_j^{p*}} = 0$; $i \in \{c, a, v\}$, where $\overline{P_j^{p*}}$ are the mean prices (Diewert and Wales, 1987, p. 54).

The shadow prices of labor can be obtained by

$$P_{Ln}^* = \frac{\partial \Pi(p_n^{p*}, r_n, X_{Ln})}{\partial X_{Ln}} = \Pi^L(p_n^{p*}, r_n) \tag{8.22}$$

In the second step we analyse the labor supply and labor demand of the households. It follows from our theoretical model that if a household participates as a seller in the off-farm labor market, the internal wage rate P_L^* equals the effective marginal off-farm labor wage P_L^s, and if a household participates as a buyer in the on-farm labor market the internal wage rate P_L^* equals the effective marginal on-farm labor wage P_L^h. Thus, the corresponding labor wage functions could be econometrically estimated based on equations (8.5) and (8.7):

$$P_L^* = \beta_0^s + X_L^s \beta_1^s + z_p^s \beta^s + \nu^s \tag{8.23}$$

$$P_L^* = \beta_0^h + X_L^h \beta_1^h + z_p^h \beta^h + \nu^h \tag{8.24}$$

According to the specification of the off-farm labor wage function, z_p^s includes factors that explain PTCs of supplying labor (z_v^s), the average skill level of a farm household (z_L^s), and the average regional wage level $(\overline{P}_L)$. Analogously, z_p^h includes factors explaining PTCs of hiring labor (z_v^h), the average skill of hired on-farm labor (z_L^h), and, again, the average regional wage level $(\overline{P}_L)$. VTCs and observable heterogeneity of skills across farm family members or across hired on-farm workers are captured by the coefficients β_1^s and β_1^h, respectively. In particular, a negative β_1^s implies increasing VTCs or heterogeneity of skills among family members, while a positive β_1^h implies increasing VTCs or decreasing productivity of hired workers. Thus, given the specification in (8.23) and (8.24), we cannot differentiate between increasing transaction costs and heterogeneity. However, our main purpose is to model and test for non-separability and this can be easily done via the β_1 parameters based on the specification above.

As labor supply and labor demand depend on the household's decision to participate in the off-farm and on-farm labor market, estimating the shadow price equations (8.23) and (8.24) might be plagued by a sample selection bias (Greene, 2003). Hence, we estimate these equations by applying a two-stage Heckman sample selection model to take non-random sample selection into account (Heckman, 1976). However, since X_L^s and X_L^h are endogenous, estimating these equations based on empirically observed labor supply and demand might be additionally plagued by an endogeneity bias. To account for possible endogeneity biases in general, we can apply a two-stage least squares (2SLS) estimation. Moreover, as X_L^s and X_L^h depend on labor market participation, we have to apply a switching regression approach at the first stage of the 2SLS estimation to control for potential selectivity biases. The specific estimation procedure we apply can be interpreted as a straightforward extension of the two-stage probit method for simultaneous equation models with selectivity suggested by Lee et al. (1980).

In the following section we demonstrate this procedure for the estimation of the labor supply function, where it applies analogously to the estimation of the labor demand function.

First, the two dichotomous variables Y^s and Y^h indicate the regime into which the observation falls.

$$Y^s = \begin{cases} 1 & \text{if} \quad Y^{s*} > 0 \\ 0 & \text{if} \quad Y^{s*} \leq 0 \end{cases} \tag{8.25}$$

$$Y^h = \begin{cases} 1 & \text{if} \quad Y^{h*} > 0 \\ 0 & \text{if} \quad Y^{h*} \leq 0 \end{cases} \tag{8.26}$$

where Y^{s*} and Y^{h*} are not observable, but indicate whether a household supplies or demands labor, respectively. Therefore, we estimate a bivariate probit model while considering that households decide simultaneously on labor supply and labor demand:

$$Y^{s*} = z\gamma^s + \varepsilon^s \tag{8.27}$$

$$Y^{h*} = z\gamma^h + \varepsilon^h \tag{8.28}$$

where $z = (\iota', z_p^{\pi\prime}, z_p^{u\prime}, z_p^{s\prime}, z_p^{h\prime}, z_f^{s\prime}, z_f^{h\prime})'$ are factors influencing labor demand and supply decisions, and ι is a vector of ones. According to our theoretical model, labor demand and supply decisions are determined by variables that influence the shadow price from the production side (z_p^{π}), variables that influence the shadow price from the consumption side (z_p^{u}), variables that explain the effective wage rate of supplied off-farm (z_p^{s}) and hired on-farm labor (z_p^{h}), and fixed transaction costs occurring for supplying (z_f^{s}) and demanding labor (z_f^{h}). Further, we assume that the residuals ν^s, ε^s, and ε^h have a trivariate normal distribution with mean 0 and covariance matrix

$$E\left[\begin{pmatrix}\nu^s\\ \varepsilon^s\\ \varepsilon^h\end{pmatrix}\begin{pmatrix}\nu^s\\ \varepsilon^s\\ \varepsilon^h\end{pmatrix}'\right] = \begin{pmatrix}\sigma & \sigma^s & \sigma^h\\ & 1 & \rho\\ & & 1\end{pmatrix} \tag{8.29}$$

Considering that we can estimate the labor supply function (8.23) only for households that participate in the market, we have

$$E[\nu^s | Y^{s*} > 0] = \sigma^s \frac{\phi(z\gamma^s)}{\Phi(z\gamma^s)} \tag{8.30}$$

where $\phi()$ and $\Phi()$ denote the probability density (pdf) and cumulative distribution (cdf) function of a standard normal distribution, respectively. Thus, we can correct for the sample selection bias by estimating

$$P_L^* = \beta_0^s + X_L^s\beta_1^s + z_p^s\beta^s + \sigma^s \frac{\phi(z\widehat{\gamma}^s)}{\Phi(z\widehat{\gamma}^s)} + \widetilde{\nu}^s \tag{8.31}$$

To account for the endogeneity of X_L^s, we apply a 2SLS estimation by substituting the fitted values $\widehat{X}_L^s$ for the observed values X_L^s in (8.31). For those households that simultaneously supply and demand labor, the optimal labor allocation depends on equations (8.11) to (8.14), while for those households that only supply labor, the optimal labor allocation depends only on (8.11), (8.12), and (8.13). Therefore, the impact of exogenous variables on the amount of supplied labor X_L^s corresponds to two different regimes. Hence, the first stage of the 2SLS estimation corresponds to the following switching regression model:

$$X_L^s = \delta^b z_x^b + \xi^b \quad \text{if} \quad Y^{s*} > 0 \ \wedge \ Y^{h*} > 0 \tag{8.32}$$

$$X_L^s = \delta^s z_x^s + \xi^s \quad \text{if} \quad Y^{s*} > 0 \ \wedge \ Y^{h*} \leq 0 \tag{8.33}$$

where $z_x^b = (\iota', z_p^{\pi\prime}, z_p^{u\prime}, z_p^{s\prime}, z_p^{h\prime})'$ and $z_x^s = (\iota', z_p^{\pi\prime}, z_p^{u\prime}, z_p^{s\prime})'$.

The assumption that each of the disturbance terms ξ^b and ξ^s is trivariate normal distributed with the error terms ϵ^s and ϵ^h results in the following conditional expectations:

$$E\left[\xi^b \left| Y^{s*} > 0 \ \wedge \ Y^{h*} > 0\right.\right] = \sigma^{bs}\lambda^{bs} + \sigma^{bh}\lambda^{bh} \tag{8.34}$$
$$E\left[\xi^s \left| Y^{s*} > 0 \ \wedge \ Y^{h*} < 0\right.\right] = \sigma^{ss}\lambda^{ss} + \sigma^{sh}\lambda^{sh} \tag{8.35}$$

where $\sigma^{bs} = cov(\xi^b, \varepsilon^s)$, $\sigma^{bh} = cov(\xi^b, \varepsilon^h)$, $\sigma^{ss} = cov(\xi^s, \varepsilon^s)$, $\sigma^{sh} = cov(\xi^s, \varepsilon^h)$, and

$$\lambda^{bs} = \frac{\phi(z\gamma^s)\Phi\left(\frac{z\gamma^h - \rho z\gamma^s}{\sqrt{1-\rho^2}}\right)}{\Phi_2(z\gamma^s, z\gamma^h)} \tag{8.36}$$
$$\lambda^{bh} = \frac{\phi(z\gamma^h)\Phi\left(\frac{z\gamma^s - \rho z\gamma^h}{\sqrt{1-\rho^2}}\right)}{\Phi_2(z\gamma^s, z\gamma^h)} \tag{8.37}$$
$$\lambda^{ss} = \frac{\phi(z\gamma^s)\Phi\left(\frac{-z\gamma^h + \rho z\gamma^s}{\sqrt{1-\rho^2}}\right)}{\Phi_2^*(z\gamma^s, -z\gamma^h)} \tag{8.38}$$
$$\lambda^{sh} = -\frac{\phi(z\gamma^h)\Phi\left(\frac{z\gamma^s - \rho z\gamma^h}{\sqrt{1-\rho^2}}\right)}{\Phi_2^*(z\gamma^s, -z\gamma^h)} \tag{8.39}$$

where Φ_2 and Φ_2^* are bivariate standard normal distributions with correlations ρ and $-\rho$, respectively.[9]

Given the conditional expectation values above, we can correct for sample selection biases following a straightforward extension of the two-stage method suggested by Lee et al. (1980).[10] In particular, we estimate in the first stage of the 2SLS estimation:

$$X_L^s = \delta^b z_x^b + \sigma^{bs}\widehat{\lambda}^{bs} + \sigma^{bh}\widehat{\lambda}^{bh} + \widetilde{\xi}^b \quad \text{if} \quad Y^{s*} > 0 \ \wedge \ Y^{h*} > 0 \tag{8.40}$$
$$X_L^s = \delta^s z_x^s + \sigma^{ss}\widehat{\lambda}^{ss} + \sigma^{sh}\widehat{\lambda}^{sh} + \widetilde{\xi}^s \quad \text{if} \quad Y^{s*} > 0 \ \wedge \ Y^{h*} \leq 0, \tag{8.41}$$

[9] A detailed derivation of the selectivity terms is available from the authors upon request.

[10] We thank Awudu Abdulai who pointed out that Saha et al. (1994) analyse a similar sample selection problem. In particular, they suggest an extended Heckman procedure, which is also applied by Abdulai et al. (2005). Although we have been stimulated by their work, we actually derived different selectivity terms. To compare our results with the results of Saha et al. (1994) we calculated the conditional expectation values numerically. Our formula perfectly fits the numerical calculation, while the formula of Saha et al. (1994) did not.

where $\widehat{\lambda}$ indicates the estimated λs calculated from the estimated parameters of the bivariate probit function. Furthermore, we calculate $\widehat{X}_L^s$ by:

$$\widehat{X}_L^s = \begin{cases} \widehat{\delta}^b z_x^b + \widehat{\sigma}^{bs}\,\widehat{\lambda}^{bs} + \widehat{\sigma}^{bh}\,\widehat{\lambda}^{bh} & \text{if} \quad Y^{s*} > 0 \,\wedge\, Y^{h*} > 0 \\ \widehat{\delta}^s z_x^s + \widehat{\sigma}^{ss}\,\widehat{\lambda}^{ss} + \widehat{\sigma}^{sh}\,\widehat{\lambda}^{sh} & \text{if} \quad Y^{s*} > 0 \,\wedge\, Y^{h*} \leq 0 \end{cases} \tag{8.42}$$

The covariance matrix of the second stage coefficients is calculated by the formula published in Lee et al. (1980).

From the estimations described above and neglecting FTC, the net off-farm labor revenue function ($f(X_L^s)$) and the effective cost function for hired labor ($g(X_L^h)$) can be obtained by:

$$f(X_L^s) - TC_f^s = \int_0^{X_L^s} \left(\widehat{\beta}_0^s + \widehat{\beta}^s z_p^s + \widehat{\beta}_1^s X_L^s\right) \mathrm{d}\, X_L^s = \left(\widehat{\beta}_0^s + \widehat{\beta}^s z_p^s\right) X_L^s + \frac{1}{2}\widehat{\beta}_1^s X_L^{s\,2} \tag{8.43}$$

$$g(X_L^h) - TC_f^h = \int_0^{X_L^h} \left(\widehat{\beta}_0^h + \widehat{\beta}^h z_p^h + \widehat{\beta}_1^h X_L^h\right) \mathrm{d}\, X_L^h = \left(\widehat{\beta}_0^h + \widehat{\beta}^h z_p^h\right) X_L^h + \frac{1}{2}\widehat{\beta}_1^h X_L^{h\,2} \tag{8.44}$$

Our estimation strategy does not allow an estimation of FTC, because TC_f^s and TC_f^h cannot be identified. However, because we are only interested in the impact of imperfect labor markets on tax-induced farm household reactions, we do not need to identify fixed transaction costs at this stage and we let them be captured by variable E (exogenous transfers).

A concave labor revenue function (f) requires a negative parameter β_1^s, and a convex cost function for hired labor (g) requires a positive parameter β_1^h. Note however, that although a convex and concave labor cost and revenue function make the FHM equilibrium analytically more traceable, it is not a necessary condition for its existence. If the estimated parameters β_1^s and β_1^h have the right sign, the standard errors of these parameters returned by our extended Heckman estimation allow us to use a simple t-test to determine if the labor revenue and labor cost functions are significantly concave and convex, respectively. However, it is important to note that the formula given by Lee et al. (1980) must be used to calculate unbiased standard errors.

In the third step we again estimate the netput equations of an SNQ profit function but this time we consider labor as a variable input. Thus, we have four netputs (X_i; $i \in \{c, a, v, L\}$) and two quasi-fixed inputs (R_i; $i \in \{g, k\}$):

$$\begin{aligned} X_{in}(p_n^{p*}, r_n) = {} & \alpha_i + w_n^{-1} \sum_{j \in \{c,a,v,L\}} \beta_{ij} P_{jn}^{p*} - \frac{1}{2}\theta_i w_n^{-2} \sum_{j \in \{c,a,v,L\}} \sum_{k \in \{c,a,v,L\}} \beta_{jk} P_{jn}^{p*} P_{kn}^{p*} \\ & + \sum_{j \in \{g,k\}} \delta_{ij} R_{jn} + \frac{1}{2}\theta_i \sum_{j \in \{g,k\}} \sum_{k \in \{g,k\}} \gamma_{jk} R_{jn} R_{kn} \end{aligned} \tag{8.45}$$

In the final step we estimate the household's consumption decisions via an AIDS consumer demand system consisting of three commodity groups: purchased commodities (C_m), self-produced consumption goods (C_a), and leisure (C_L). We use the following specification (Deaton and Muellbauer, 1980):

$$W_{in} = \alpha_i + \sum_{j \in \{m,a,L\}} \gamma_{ij} \ln P^{c*}_{jn} + \beta_i \ln \frac{Y_n}{\wp_n} \tag{8.46}$$

$$\ln \wp_n = \alpha_0 + \sum_{i \in \{m,a,L\}} \alpha_i \ln P^{c*}_{in} + \frac{1}{2} \sum_{i \in \{m,a,L\}} \sum_{j \in \{m,a,L\}} \gamma_{ij} \ln P^{c*}_{in} \ln P^{c*}_{jn} \tag{8.47}$$

Here, $W_{in} = P^{c*}_{in} C_{in} / Y_n$; $i \in \{m, a, L\}$ are the budget shares, where Y_n indicates full income. $\wp_n$ is the translog consumer price index, P^{c*}_{in}; $i \in \{m, a, L\}$ indicates the consumer prices, and α, β, and γ are the parameters to be estimated.

The estimation of the second-step profit function and the demand system might be biased by measurement errors and endogeneity problems. First, the price of labor or leisure (P_L) is not exactly measured but estimated from the results of the first-step profit function (8.22). Second, full income (Y) used as an explanatory variable in the demand system might be endogenous and depends on the estimated price of leisure. To avoid these estimation biases we apply a three-stage least squares (3SLS) estimation, where we again use the above specified variables z as instrumental variables for P^{p*}_L, P^{c*}_L and Y.

8.5 Data and Empirical Results

Data used for the estimations are based on an accounting survey of 202 agricultural households in several regions around Poznan (Mid-West Poland) in 1994. The data were collected by the Institute for Agriculture and Food Industries in Warsaw (IERiGZ, 1995). Additional regional data used in this analysis are taken from Glowny Urzad Statystyczny (1996) and Zawadzki (1994). Characteristics of the farm sample are summarized in table 8.A1 in the appendix. Additionally, sample characteristics of different labor market regimes are presented in table 8.A2 in the appendix.

On the production side, market goods (X_c) consist of all crop products, while animal products are considered as (possibly) home-consumed goods (X_a). All relevant variable inputs of the farms are subsumed in netput X_v. Labor (X_L) includes both family (X^f_L) and hired labor (X^h_L). Land (R_g) and capital (R_k) are considered as quasi-fixed factors. On the consumption side, C_m includes all purchased consumption goods. The self-produced goods (C_a) correspond conceptually to the home-consumed animal products (X_a). The amount of leisure (C_L) is determined by calculating the yearly available time (T_L) of households[11] minus on-farm (X_L^f) and off-farm (X_L^s) family labor.

The variables z_p^π influencing the shadow price from the production side include resource endowments (R_g, R_k) as well as variable output and input prices. The variables z_p^u influencing the shadow price from the consumption side include household composition and consumer prices. Household composition is measured by the number of family members up to 14 years (N_k), between 15 and 60 years (N_w), and older than 60 years (N_o). Moreover, z_p^u includes sex (D_f), age (A_h), and age squared (A_h^2) of the head of the household.

The explanatory variables for variable transaction costs of supplying labor (z_v^s) include the regional density of the railroad net (W_r), the regional density of the road net (W_s), the regional number of telephones per household (W_t), the regional unemployment rate (W_u), and the number of cars owned by the household (N_c). The variables explaining the average off-farm skill level of farm households (z_L^s) include the number of family members that are of working age (N_w), the number of family members older than 60 years (N_o), and the average level of human capital. Since we had no data on education, we follow Vakis et al. (2003) and interpret sex (D_f), age (A_h), and age squared (A_h^2) of the head of household as indicators of average human capital.

The sample contains two farms that do not produce any animal products, which are removed to provide a more homogeneous sample and to avoid an imputation of the unknown prices of animal products.

All estimations and calculations are carried out by the (free) statistical software 'R' (R Development Core Team, 2005, see also http://www.r-project.org), using the add-on packages 'micEcon' (Henningsen and Toomet, 2005), 'systemfit' (Hamann and Henningsen, 2005) and 'VGAM' (Yee and Wild, 1996).

8.5.1 Estimation results

In the first step the three netput equations of the SNQ profit function (8.21) are estimated; the results are presented in appendix table 8.A3. The R^2 values are 0.71, 0.29, and 0.69 for X_c, X_a, and X_v, respectively, and 14 out of 18 coefficients are significantly different from zero.

The homogeneity and symmetry conditions are imposed in this estimation. While monotonicity is fulfilled at all observations, the estimated profit function is not convex in prices.

[11]It is assumed that each household member between 15 years and 60 years has 10 hours per day and each household member older than 60 years has 5 hours per day available for work and/or leisure. The annual available time of the household is calculated by multiplying the total hours per day of all household members by 365.

We first tried to impose convexity by a non-linear estimation using the Cholesky decomposition (Lau, 1978b). Because the estimation of the restricted non-linear netput equations did not converge, we chose a new procedure suggested by Koebel et al. (2000, 2003) based on the minimum distance and asymptotic least squares estimation (Gourieroux et al., 1985; Kodde et al., 1990), and asymptotically equivalent to a (successful) non-linear estimation with convexity imposed. First, the estimation results of the unrestricted (linear) netput equations are used to calculate the Hessian matrix of the unrestricted profit function. Second, the weighted difference between this unrestricted and a restricted Hessian matrix is minimized. Finally, restricted coefficients are identified by an asymptotic least squares (ALS) framework.

The weighting matrix for the minimization of the difference between the unrestricted and the restricted Hessian matrix is the inverse of the variance-covariance matrix of the Hessian matrix, which can be derived from the variance-covariance matrix of the estimated coefficients. We use the Cholesky factorization to restrict the Hessian matrix to be positive semi-definite.[12]

The parameter estimates and R^2 values of the restricted profit function are presented in appendix table 8.A4. The R^2 values are almost identical to the unrestricted model. This shows that the data do not unreasonably contradict the convexity constraint. The homogeneity and symmetry conditions are not affected by the imposition of convexity, and monotonicity is still fulfilled for all observations. Thus, the estimated profit function fully complies with microeconomic theory.

The shadow prices of labor calculated from the restricted profit function have reasonable values for all but one farm household. This household has a negative shadow price and is therefore removed from the sample. Hence, the sample used for the further analysis includes 199 farm households.

The labor supply and demand of the households are analyzed in the second step. The results of the estimation of labor demand and supply functions are summarized in table 8.2. First, results of the bivariate probit estimation show that labor demand and supply decisions are not significantly correlated in the sample (ρ is not significantly different from zero). Moreover, the expected signs have been revealed for most of the factors influencing labor market participation, although the impact of many factors is not statistically significant. Since labor

[12] To retain convexity of the SNQ profit function, it is sufficient to minimize the difference between the estimated (unrestricted) β-coefficients and the (linearly independent) values of a restricted β-coefficient matrix (Koebel, 1998). This procedure only allows us to adjust the β-coefficients, while the approach of Koebel et al. (2000, 2003) adjusts *all* coefficients. Thus, the fit of the constrained model is much better, due to the flexibility of the other coefficients. Both approaches 'produce' the same βs.

Table 8.2: Estimated coefficients of labor market analysis

	Labor supply		Labor demand	
Regressor	1st step: probit	2nd step: 2SLS	1st step: probit	2nd step: 2SLS
Constant	2.900	36.485	2.917	1.134
X_L^s/T_L		-80.612 ***		
X_L^h				0.047 ***
N_k	0.113		0.088	
N_w	0.154 *	-4.114 **	-0.386 ***	
N_o	0.018	-4.500 *	-0.306 **	
D_f	0.391	-9.680	-0.195	
A_h	0.006	2.247 *	-0.118 *	
A_h^2	$-1.6 \cdot 10^{-4}$	-0.025 *	$1.3 \cdot 10^{-3}$ *	
R_g	-0.013		0.025	
R_k	$-2.2 \cdot 10^{-8}$		$1.3 \cdot 10^{-6}$ ***	
P_c	2.520		4.462 *	
P_a	0.225		-0.311	
P_v	-2.955		-3.090	
N_c	0.145	-2.879	-0.142	
$\overline{P}_L$	-0.015	0.042	-0.000	-0.075
W_u	-0.008	-0.226	-0.041	1.295
W_s	-0.017	-0.235	-0.013	0.380
W_r	-0.097 *	2.410 *	0.044	2.813 **
W_t	-0.003	-0.401	0.003	-0.528
IMR Supply		-8.577		
IMR Demand				-18.358 **
ρ	-0.140		-0.140	
R^2		0.306		0.388

Notes: IMR = inverse Mill's ratio.
'*', '**', and '***' denote statistical significance at the 10%, 5%, and 1% level, respectively.

market participation is not the main focus in this paper we do not further comment on the estimation results in detail.[13]

However, the signs and the statistical significance of the estimated parameters of the amount of supplied and demanded labor are of particular interest because these parameters determine the curvature properties of the labor revenue and cost function, respectively.

[13]A detailed interpretation of these estimation results is given in Henning and Henningsen (2005a).

As can be seen from table 8.2, the parameter for supplied labor[14] is significantly negative, while the parameter for the demanded labor amount is significantly positive. Hence, the estimated labor revenue and cost function is found to be significantly concave and convex, respectively, indicating significant imperfections in terms of increasing VTCs or heterogeneity for both labor markets.

In the third step we estimate the four netput equations of the SNQ profit function (8.45). The parameter estimates and R^2 values are presented in appendix table 8.A5. The R^2 values are 0.75, 0.49, 0.82, and 0.28 for X_c, X_a, X_v, and X_L, respectively, and 17 out of 25 parameters are significantly different from zero. Again, the homogeneity and symmetry conditions are imposed in the estimation. Monotonicity is fulfilled at 98.0% of the observations, but the estimated profit function is not convex in prices. Hence, convexity is enforced using the same method as for the profit function in the first step. The parameter estimates and R^2 values of the restricted profit function are presented in appendix table 8.A6. Again, the R^2 values are almost identical to the unrestricted model, showing that the data do not unreasonably contradict the convexity constraint. Since the homogeneity and symmetry conditions are maintained in the imposition of convexity, and the restricted profit function fulfills monotonicity at 97.0% of the observations, microeconomic theory is satisfied for almost the complete sample.

In the final step the budget share equations of the Almost Ideal Demand System (8.46) are estimated by the Iterated Linear Least Squares Estimator (ILLE) of Blundell and Robin (1999). The estimation results are presented in appendix table 8.A7. The R^2 values are 0.41, 0.59, and 0.50 for W_m, W_a, and W_L, respectively, and 9 out of 12 parameters are significantly different from zero. The adding-up, homogeneity, and symmetry conditions are imposed on the estimated parameters. Monotonicity is fulfilled at 99.5% of the observations and concavity is fulfilled at 88.4% of the observations. Thus, the estimated demand system meets the conditions derived from microeconomic demand theory in a very large range of the sample.

8.5.2 Tax Elasticities

The main focus of our empirical analysis is the quantitative assessment and comparison of farm households' responses to various tax policies assuming perfect, imperfect, and missing

[14] To normalize for different household sizes, we use the share of off-farm labor in total labor endowment (X_L^s/T_L) as an explanatory variable in the off-farm labor wage equation. This normalization is based on the assumption that the share of skilled and unskilled labor in the total family does not significantly vary with the family size. Using the absolute amount of off-farm labor supply instead does not change the main results, i.e. significant and negative impact on the effective off-farm wage rate.

markets. We use a specific non-standard concept of elasticities to facilitate both calculation and comparison among different institutional settings. Tax elasticities are formally calculated as follows:

$$\frac{\partial \ln Q}{\partial \tau_j} = \left.\frac{\partial \ln Q}{\partial \tau_j}\right|_{P_L^*=\text{const.}} + \frac{\partial \ln Q}{\partial \ln P_L^*}\frac{\partial \ln P_L^*}{\partial \tau_j} \tag{8.48}$$

where $Q \in \{C_i, X_i, X_l^s, X_l^h\}$ are the decision variables and $\tau_j, j \in \{y, w, ms, v, r, vat\}$ denote the tax parameters. As regards content the calculated tax elasticities indicate the relative change of a decision variable with respect to the change of the analyzed tax parameter by one percentage *point*. Technically, these have been calculated based on the relevant price and income elasticities (see appendix tables 8.A8 and 8.A9) as well as the labor market reactions, which are all based on the estimated parameters and calculated using the sample mean values of the relevant variables.[15]

Table 8.3 gives an overview of estimated tax elasticities assuming perfect, imperfect, and missing labor markets. Regarding non-standard production effects of income and consumption taxes, estimated tax elasticities partly confirm standard presumptions: even for imperfect markets standard tax policies have almost no impact on farm production, while agricultural tax policies, namely market surplus and input taxes, induce considerable production adjustment when compared to corresponding own price elasticities. However, a different picture results when assuming that labor markets are totally missing. In this case, standard taxes imply considerable production adjustments that are even partly higher when compared to corresponding impacts of market surplus and input taxes. Furthermore, a land tax implies considerable adjustments of farm production when markets are imperfect or missing, contradicting standard presumptions in public finance theory. These adjustments mainly occur for market goods (X_c) and on-farm labor (X_L). Interestingly, all non-standard effects of income, consumption, and land taxes on farm production are positive, i.e. increased taxes imply increasing farm production. According to our theoretical analysis this mainly follows from the fact that increased taxes induce significant decreases of the endogenous labor wage, which in turn stimulates farm production. The lower the potential of a farm household to adapt its labor-leisure choice via labor markets the higher is c.p. the induced shadow price effect. Therefore, non-standard tax impacts are generally higher for missing labor markets when compared to existing but imperfect labor markets.

Regarding classical surrogates of the income tax applied to peasant economy, like market surplus and input taxes, adjustment patterns (signs of tax elasticities) are similar when perfect and imperfect markets are compared. In contrast, assuming missing markets implies

[15] A detailed derivation of the tax elasticities is available from the authors upon request.

Table 8.3: Tax elasticities

	τ_y	τ_w	τ_{ms}	τ_v	τ_g	τ_{vat}
Separable model: Perfect labour market						
X_c	0.00	0.36	-0.93	-0.57	0.00	0.00
X_a	0.00	0.12	-0.85	-0.73	0.00	0.00
X_v	0.00	0.00	-1.08	-1.08	0.00	0.00
X_L	0.00	0.51	-0.51	-0.00	0.00	0.00
C_m	-0.55	-0.42	-0.46	-0.14	-0.20	-0.67
C_a	-0.31	-0.15	0.38	-0.18	-0.26	0.50
C_L	-0.26	0.16	-1.05	-0.46	-0.66	-0.19
X_L^{sn}	8.38	-13.13	41.35	14.83	20.93	6.16
Non-separable model: Imperfect labour market						
X_c	0.07	0.24	-0.60	-0.39	0.17	0.05
X_a	0.02	0.08	-0.74	-0.68	0.05	0.02
X_v	0.00	0.00	-1.08	-1.08	0.00	0.00
X_L	0.09	0.34	-0.05	0.25	0.24	0.07
C_m	-0.62	-0.29	-0.83	-0.41	-0.39	-0.72
C_a	-0.34	-0.12	0.26	-0.34	-0.32	0.48
C_L	-0.23	0.05	-0.87	-0.60	-0.57	-0.17
X_L^{sn}	5.72	-6.71	28.23	15.02	14.29	4.21
X_L^{s}	2.31	-4.23	11.41	6.07	5.78	1.70
X_L^{h}	-0.60	-2.13	-2.95	-1.57	-1.49	-0.44
X_L^{f}	0.15	0.54	0.19	0.39	0.38	0.11
P_L	-1.19	-0.66	-0.92	-0.49	-0.46	-0.14
Non-separable model: Missing labour market						
X_c	0.22	0.00	0.14	0.00	0.54	0.16
X_a	0.07	0.00	-0.51	-0.55	0.17	0.05
X_v	0.00	0.00	-1.08	-1.08	0.00	0.00
X_L	0.30	0.00	0.99	0.80	0.76	0.22
C_m	-0.79	0.00	-1.66	-0.85	-0.81	-0.85
C_a	-0.39	0.00	-0.02	-0.48	-0.46	0.44
C_L	-0.15	0.00	-0.49	-0.39	-0.37	-0.11
P_L	-1.60	0.00	-2.94	-1.56	-1.49	-0.44

different adjustment patterns, especially on the production side. When labor markets are perfect a market surplus and input tax reduce farm production due to lower decision prices of outputs. However, when labor markets are imperfect these effects are counterbalanced by induced decreases of the endogenous labor wage. As long as the imperfection is not too strong, the main adjustment patterns sustain, although effects are lower in absolute terms. When markets are totally missing, counterbalanced effects partly overcompensate direct effects of increased taxes. For instance, production of market goods and on-farm labor are increased due to these taxes.

The wage tax is a special case. As we have already explained above, farm responses to a wage tax are clearly driven by market participation. Thus, since we assume that households participate in off-farm markets under both perfect and imperfect markets, adjustment patterns are relatively similar for these cases. In contrast, logically a wage tax has no impact on farm production and consumption if households do not participate in off-farm labor markets, which is the case when labor markets are missing.

Regarding adjustments of consumption as well as net-labor supply, adjustment patterns are relatively similar for all labor market conditions. In particular, consumption is reduced due to reduced household income, as is implied by increased taxes. Moreover, net-labor supply is remarkably decreased for the wage tax, while all other tax policies imply a remarkable increase. Interestingly, in absolute terms tax responses are generally higher under imperfect markets when compared to perfect markets for standard taxes as well as for a land tax, while vice-versa these are generally higher under perfect markets for a market surplus, input, and wage tax. Obviously, market imperfections reduce the adaption capacity of farm households and, therefore, generally limit labor and leisure adjustments to tax policies. In some cases (e.g. standard taxes) limited labor adjustments amplify adjustments of market goods, while in other cases (e.g. agricultural taxes) it reduces adjustments of market goods.

Considering our empirical results, we partly have to weaken our conclusions drawn from the theoretical analysis. At least for the Polish case, standard taxes have almost no impact on farm production, even if labor markets are imperfect but not totally missing. Thus, assuming that no ex ante distortions exist, these taxes still seem to be superior to specific agricultural taxes from an economic welfare point of view, even in the case of imperfect market institutions. Moreover, for most of the analyzed taxes both signs and also absolute values of tax elasticities are rather similar when perfect and imperfect markets are compared. Thus, one might conclude that our empirical results finally confirm standard presumptions of public finance theory, even when market imperfection is taken into account. But, given the empirical results above, this certainly would be a far too fast conclusion for at least three reasons.

First, our empirical results for missing markets clearly contradict standard presumptions. Hence, income and consumption taxes have considerable impacts on farm production and agricultural taxes have deviating adjustment patterns. Second, estimated tax elasticities for the land tax imply remarkable production adjustments for both imperfect and missing labor markets. Third, when market institutions are imperfect or missing, optimal taxation analysis has to take ex ante distortions into consideration. Therefore, any differences in quantitative farm responses between the various tax policies by no means allow a direct conclusion regarding the priority of tax policies from a welfare economic perspective.

Thus, overall, our empirical results clearly underline that a comprehensive analysis of taxation of farm households has to be based on a quantitative assessment of farm responses. Moreover, it has to take the organisation of rural markets into account, because the rural market organisation has a significant impact on both positive comparative static analysis and normative evaluation of tax policies.

8.6 Conclusion

Stimulated by an ongoing discussion on the optimal design of peasant tax systems, the paper presents a comprehensive analysis of standard and selected agricultural taxes within an interdependent farm household model (FHM). In contrast to former interdependent FHM approaches, our modeling strategy allows a consistent theoretical and empirical analysis of standard and agricultural tax policies for perfect, imperfect, and missing labor markets. In particular, we overcome major shortcomings in the theoretical conception as well as serious selectivity and endogeneity problems in econometric estimation of labor supply and demand functions, which are inherent in earlier work that take labor market imperfections into account.

Major results of our theoretical and empirical analyses are: (i) Comparative static analysis implies that neither standard nor land taxes are theoretically superior when compared to agricultural tax instruments. This supports the results of recent studies. (ii) However, empirical estimation with individual household data from Poland partly confirms standard presumptions in public finance literature. Even when labor markets are imperfect, income and consumption taxes imply negligible production adjustments, while market surplus and input taxes cause remarkable production responses. (iii) In contrast, in case of totally missing markets, standard taxes provoke notable changes of farm production and a land tax implies considerable production adjustments for all labor market conditions.

Thus, we conclude that our results clearly underline that (1) the rural market conditions have a significant impact on both positive comparative static analysis and normative evaluation of tax policies. (2) Given imperfect formal market institutions, no simple rule of thumb seems to exist, making optimal agricultural tax policy design even more complex in transition and developing countries. (3) A comprehensive analysis of taxation of farm households has to be based on a quantitative assessment of farm responses to tax policies.

Bibliography

Abdulai, A., Monnin, P., Gerber, J., 2005. Output uncertainty, information acquisition and the adoption of new technologies. Unpublished Working Paper, University of Kiel, Germany.

Ahmad, E., Stern, N., 1991. The Theory and Practice of Tax Reform in Developing Countries. Cambridge University Press, Cambridge.

Atkinson, A. B., 1987. The theory of tax design for developing countries. In: Newbery, D. M. G., Stern, N. H. (Eds.), The Theory of Taxation for Developing Countries, Oxford University Press, New York, pp. 387–404.

Besley, T., 1993. Targeting taxes and transfers: Administrative costs and policy design in developing economies. In: Hoff, K., Braverman, A., Stiglitz, J. E. (Eds.), The Economics of Rural Organization, Oxford University Press, New York, pp. 374–405.

Blundell, R. W., Robin, J. M., 1999. Estimation in large and disaggregated demand systems: An estimator for conditionally linear systems. Journal of Applied Econometrics 14, 209–232.

Burgess, R., Stern, N., 1993. Taxation and development. The Journal of Economic Literature 31 (2), 762–830.

Chambers, R. G., Lopez, R. E., 1987. Tax policies and the financially constrained farm household. American Journal of Agricultural Economics 69 (2), 369–377.

de Janvry, A., Fafchamps, M., Sadoulet, E., 1991. Peasant household behaviour with missing markets: some paradoxes explained. The Economic Journal 101 (409), 1400–1417.

Deaton, A., Muellbauer, J., 1980. An Almost Ideal Demand System. The American Economic Review 70, 312–326.

Diamond, P. A., Mirrless, J., 1971. Optimal taxation and public production. part i: Production and efficiency. American Economic Review 61 (1), 8–27.

Diewert, W. E., Wales, T. J., 1987. Flexible functional forms and global curvature conditions. Econometrica 55 (1), 43–68.

Diewert, W. E., Wales, T. J., 1992. Quadratic spline models for producer's supply and demand functions. International Economic Review 33, 705–722.

Fafchamps, M., 1992. Cash crop production, food price volatility, and rural market integration in the third world. American Journal of Agricultural Economics 74, 90–99.

Finkelshtain, I., Chalfant, J. A., 1991. Marketed surplus under risk: do peasants agree with sandmo? American Journal of Agricultural Economics 73, 557–567.

Glauben, T., 2000. Besteuerung im Agrarsektor: eine theoretische und empirische Analyse auf der Grundlage von Unternehmens-Haushalts-Modellen. No. 21 in Agrarökonomische Studien, Vauk, Kiel.

Glauben, T., Henning, C. H. C. A., Henningsen, A., 2003. Farm household decisions under various tax policies: Comparative static results and evidence from household data. Paper presented at the 25th International Conference of the IAAE, Durban, South Africa.

Glowny Urzad Statystyczny, 1996. Informacja o sytuacji spoleczno - gospodarczej wojewodztw. Kwartalnik Nr 1-2, Warsaw.

Gourieroux, C., Monfort, A., Trognon, A., 1985. Moindres carres asymptotiques. Annales de l'INSEE 58, 91–122.

Greene, W. H., 2003. Econometric Analysis. Prentice Hall, 5th ed.

Hamann, J. D., Henningsen, A., 2005. systemfit: Simultaneous Equation Estimation. R package version 0.7-6, http://cran.r-project.org.

Heady, C. J., Mitra, P. K., 1987. Optimal taxation and shadow pricing in a developing economy. In: Newbery, D. M. G., Stern, N. H. (Eds.), The Theory of Taxation for Developing Countries, Oxford University Press, New York, pp. 407–425.

Heckman, J., 1976. The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement 5, 475–492.

Henning, C. H. C. A., 1994. Unternehmens-Haushalts-Modelle. Duncker & Humblot, Berlin.

Henning, C. H. C. A., 1995. Theoretische Ableitung eines UHM-Ansatzes unter expliziter Berücksichtigung von unvollkommenen Märkten und Steuerpolitiken, mimeo, Mannheimer Zentrum für Europäische Sozialpolitik (MZES).

Henning, C. H. C. A., 1998. Ökonometrische Schätzung eines interdependenten UHM-Ansatzes zur Analyse von Steuerpolitik in der polnischen Landwirtschaft unter expliziter Berücksichtung unvollkommener Arbeitsmärkte, mimeo, Stanford University.

Henning, C. H. C. A., Henningsen, A., 2005a. Modeling price response of farm households in imperfect labour markets in Poland: Incorporating transaction costs and heterogeneity into a farm household approach, mimeo, University of Kiel, Germany.

Henning, C. H. C. A., Henningsen, A., 2005b. Modelling price response of farm households under imperfect labour markets: A farm household approach to family farms in poland. Paper presented at the 94th EAAE Seminar in Wye, UK, 9-10 April, http://www.eaae-wye.org.uk/papers/Henning-Heningsen.pdf.

Henningsen, A., Toomet, O., 2005. micEcon: Tools for Microeconomic Analysis and Microeconomic Modeling. R package version 0.2-1, http://cran.r-project.org.

Hirschman, A. O., 1964. Land taxes and land reform in Columbia. In: Bird, R., Oldman, O. (Eds.), Readings on Taxation in Developing Countries, Johns Hopkins University Press, Baltimore.

Hoff, K., 1993. Land taxes, output taxes, and sharecropping: Was Henry George right? In: Hoff, K., Braverman, A., Stiglitz, J. E. (Eds.), The Economics of Rural Organization, Oxford University Press, New York, pp. 332–351.

Hoff, K., Braverman, A., Stiglitz, J. E. (Eds.), 1993. The Economics of Rural Organization. Oxford University Press, New York.

Hoff, K., Stiglitz, J. E., 1993. Consequences of limited risk markets and imperfect information for the design of taxes and transfers: An overview. In: Hoff, K., Braverman, A., Stiglitz, J. E. (Eds.), The Economics of Rural Organization, Oxford University Press, New York, pp. 323–331.

IERiGZ, 1995. Indywiualne wyniki rachunkowosci rolnej gospodarstw chlopskich - Rok 1991-1994 - Makroregion VIII - Srodkowo Zachodni. Unpublished data, personally requested from IERiGZ, Warsaw, Poland.

Key, N., Sadoulet, E., de Janvry, A., 2000. Transaction costs and agricultural household supply response. American Journal of Agricultural Economics 82 (2), 245–259.

Klein, L. R., 1953. A Textbook of Econometrics. Row, Peterson and Co, New York.

Kodde, D. A., Palm, F. C., Pfann, G. A., 1990. Asymptotic least-squares estimation efficiency considerations and applications. Journal of Applied Econometrics 5, 229–243.

Koebel, B., 1998. Tests of representative firm models: Results for german manufacturing industries. Journal of Productivity Analysis 10, 251–270.

Koebel, B., Falk, M., Laisney, F., 2000. Imposing and testing curvature conditions on a Box-Cox cost function. Discussion Paper No. 00-70, ZEW, Mannheim, ftp://ftp.zew.de/pub/zew-docs/dp/dp0070.pdf.

Koebel, B., Falk, M., Laisney, F., 2003. Imposing and testing curvature conditions on a Box-Cox cost function. Journal of Business and Economic Statistics 21 (2), 319–335.

Kohli, U. R., 1993. A symmetric normalized quadratic GNP function and the US demand for imports and supply of exports. International Economic Review 34 (1), 243–255.

Lau, L. J., 1978a. Applications of profit functions. In: Fuss, M., McFadden, D. (Eds.), Production Economics: A Dual Approach to Theory and Applications, Vol 1, North-Holland, Amsterdam, pp. 133–216.

Lau, L. J., 1978b. Testing and imposing monotonicity, convexity and quasi-convexity constraints. In: Fuss, M., McFadden, D. (Eds.), Production Economics: A Dual Approach to Theory and Applications, Vol. 1, North-Holland, Amsterdam, pp. 409–453.

Lee, L., Maddala, G. S., Trost, R. P., 1980. Asymptotic covariance matrices of two-stage probit and two-stage tobit methods for simultaneous equations models with selectivity. Econometrica 48 (2), 491–503.

Lopez, R., 1984. Estimating labor supply and production decisions of self-employed farm producers. European Economic Review 24, 61–82.

Munk, K. J., 1980. Optimal taxation with some non-taxable commodities. Review of Economic Studies 47 (149), 755–766.

Newbery, D., 1987. Agricultural taxation: The main issues. In: Newbery, D., Stern, N. (Eds.), The Theory of Taxation for Developing Countries, Oxford University Press, New York, pp. 163–204.

Newbery, D., Stern, N. (Eds.), 1987. The Theory of Taxation for Developing Countries. Oxford University Press, New York.

R Development Core Team, 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Ramsey, F., 1927. A contribution to the theory of taxation. Economic Journal 37, 47–61.

Rao, J. M., 1989. Taxing agriculture: Instruments and incidence. World Development 17 (6), 809–823.

Sadoulet, E., de Janvry, A., Benjamin, C., 1998. Household behavior with imperfect labor markets. Industrial Relations 37 (1), 85–108.

Sah, R. K., Stiglitz, J. E., 1987. The taxation and pricing of agricultural and industrial goods in developing countries. In: Newbery, D., Stern, N. (Eds.), The Theory of Taxation for Developing Countries, Oxford University Press, New York.

Saha, A., Love, H. A., Schwart, R. B., 1994. Adoption of emerging technologies under output uncertainty. American Journal of Agricultural Economics 76 (4), 836–846.

Sonoda, T., Maruyama, Y., 1999. Effects of the internal wage on output supply: A structural estimation of japanese rice farmers. American Journal of Agricultural Economics 81 (1), 131–143.

Strauss, J., 1986. The theory and comparative statics of agricultural household models: A general approach. In: Singh, I., Squire, L., Strauss, J. (Eds.), Agricultural Household Models: Extensions, Applications, and Policy, Johns Hopkins University Press, Baltimore, pp. 71–91.

Vakis, R., Sadoulet, E., de Janvry, A., 2003. Measuring transactions costs from observed behavior: Market choices in Peru. http://are.berkeley.edu/~sadoulet/papers/TC-wp.pdf.

Yee, T. W., Wild, C. J., 1996. Vector generalized additive models. Journal of the Royal Statistical Society, Series B, Methodological 58, 481–493.

Zawadzki, W., 1994. Infrastruktura techniczna - glowny warunek wielofunkcyjnego rozwoju obszarow wiejskich. Tech. Rep. 365, IERiGZ, Warsaw.

Appendix tables

Table 8.A1: Characteristics of the sample

Variable	Unit	Mean	Minimum	Maximum	Std.deviation
P_cX_c	1000 PLZ	132258	10451	1189412	133724
P_aX_a	1000 PLZ	212570	2669	2526524	239835
$P_v\|X_v\|$	1000 PLZ	211960	13480	2204671	213479
P_mC_m	1000 PLZ	91469	26365	280176	42853
P_aC_a	1000 PLZ	19041	1625	41853	7606
$\|X_L\|$	hours	3686	400	9843	1717
X_L^h	hours	211	0	2085	365
X_L^s	hours	446	0	4000	876
X_L^{sn}	hours	235	-2085	4000	1002
X_L^f	hours	3475	400	9236	1705
C_L	hours	7478	23	20873	4007
R_g	ha	14.7	1.2	101.5	12.4
R_k	1000 PLZ	649191	43960	4492025	554120

Notes: Calculations are based on IERiGZ (1995). PLZ = Polish Zloty.

Table 8.A2: Characteristics of the different labor regimes

Variable	Unit	All	Suppl.+dem.	Only suppl.	Only dem.	Autarkic
number		199	57	47	61	34
$P_c X_c$	1000 PLZ	132258	157581	65883	180020	95869
$P_a X_a$	1000 PLZ	212570	220643	123997	300046	164531
$P_v \lvert X_v \rvert$	1000 PLZ	211960	232143	117552	299629	151343
$P_m C_m$	1000 PLZ	91469	105939	78012	97792	74467
$P_a C_a$	1000 PLZ	19041	18487	19245	19939	18076
$\lvert X_L \rvert$	hours	3686	3579	3372	4040	3668
X_L^h	hours	211	278	0	430	0
X_L^s	hours	446	515	1266	0	0
X_L^{sn}	hours	235	237	1266	-430	0
X_L^f	hours	3475	3301	3372	3610	3668
C_L	hours	7478	7295	8254	6473	8517
R_g	ha	14.7	16.9	9.4	18.3	11.7
R_k	1000 PLZ	649191	788881	425398	816534	424132
T_L	hours	11399	11110	12891	10082	12185
N_k		1.3	1.5	1.3	1.4	0.7
N_w		2.8	2.8	3.2	2.4	3.0
N_o		0.7	0.6	0.6	0.8	0.7
A_h	years	43	41	44	43	45
N_c		0.9	1.0	0.8	0.9	0.8

Notes: Calculations are based on IERiGZ (1995). PLZ = Polish Zloty.

Table 8.A3: Estimation results of 1st step profit function (unrestricted)

Parameter	$i = c$		$i = a$		$i = v$	
α_i	-1.72	(-0.73)	20.1	(4.31)	-17.4	(-5.14)
β_{ic}	-14.8	(-1.12)	19.8	(2.68)	-4.92	(-0.37)
β_{ia}	19.8	(2.68)	61.6	(5.76)	-81.4	(-8.04)
β_{iv}	-4.92	(-0.37)	-81.4	(-8.04)	86.3	(5.08)
δ_{ig}	6258	(11.37)	1002	(0.93)	-4306	(-5.37)
δ_{ik}	0.0829	(5.77)	0.209	(7.47)	-0.111	(-5.36)
γ_{gg}	-1157392	(-6.45)				
γ_{gk}	36.7	(7.59)				
γ_{kk}	$-1.26 \cdot 10^{-3}$	(-9.79)				
R^2	0.709		0.286		0.685	

Notes: t-values in parentheses. The standard errors of the coefficients that have not been directly estimated are calculated with the formula of Klein (1953, p. 258).

Table 8.A4: Estimation results of 1st step profit function (restricted)

Parameter	$i = c$	$i = a$	$i = v$
α_i	-2.28	20.3	-17.0
β_{ic}	3.31	14.6	-17.9
β_{ia}	14.6	64.7	-79.3
β_{iv}	-17.9	-79.3	97.3
δ_{ig}	6170	1024	-4294
δ_{ik}	0.0855	0.208	-0.110
γ_{gg}	-1149343		
γ_{gk}	36.6		
γ_{kk}	$-1.26 \cdot 10^{-3}$		
R^2	0.708	0.283	0.686

Table 8.A5: Estimation results of final profit function (unrestricted)

Parameter	$i = c$		$i = a$		$i = v$		$i = L$	
α_i	-28774	(-3.22)	32489	(2.05)	-6713	(-0.57)	-62854	(-12.61)
β_{ic}	884	(0.02)	95384	(2.76)	-61682	(-1.14)	-34586	(-4.22)
β_{ia}	95384	(2.76)	76706	(1.19)	-163015	(-2.97)	-9075	(-0.63)
β_{iv}	-61682	(-1.14)	-163015	(-2.97)	221720	(2.95)	2978	(0.24)
β_{iL}	-34586	(-4.22)	-9075	(-0.63)	2978	(0.24)	40683	(7.48)
δ_{ig}	6897	(11.68)	131	(0.12)	-6000	(-7.02)	-3158	(-8.95)
δ_{ik}	0.121	(9.02)	0.292	(12.21)	-0.166	(-9.31)	$7.41{\cdot}10^{-3}$	(0.93)
γ_{gg}	-173	(-3.55)						
γ_{gk}	$9.88{\cdot}10^{-3}$	(9.24)						
γ_{kk}	$-3.55{\cdot}10^{-7}$	(-24.28)						
R^2	0.746		0.494		0.821		0.283	

Notes: t-values in parentheses. The standard errors of the coefficients that have not been directly estimated are calculated with the formula of Klein (1953, p. 258).

Table 8.A6: Estimation results of final profit function (restricted)

Parameter	$i = c$	$i = a$	$i = v$	$i = L$
α_i	-31259	33696	-5479	-62940
β_{ic}	53069	64877	-84583	-33363
β_{ia}	64877	116801	-168354	-13324
β_{iv}	-84583	-168354	247367	5570
β_{iL}	-33363	-13324	5570	41117
δ_{ig}	6815	303	-6087	-3181
δ_{ik}	0.124	0.291	-0.167	$7.87{\cdot}10^{-3}$
γ_{gg}	-172			
γ_{gk}	$9.84{\cdot}10^{-3}$			
γ_{kk}	$-3.55{\cdot}10^{-7}$			
R^2	0.747	0.492	0.821	0.278

Table 8.A7: Estimation results of the AIDS

Parameter	$i = m$		$i = a$		$i = L$	
α_i	0.555	(9.87)	0.185	(14.79)	0.259	(4.17)
β_i	-0.170	(-9.14)	-0.031	(-7.36)	0.201	(9.94)
γ_{im}	0.034	(1.28)	0.021	(0.80)	-0.056	(-5.34)
γ_{ia}	0.021	(0.80)	0.010	(0.35)	-0.031	(-9.36)
γ_{iL}	-0.056	(-5.34)	-0.031	(-9.36)	0.086	(7.97)
R^2	0.409		0.585		0.504	

Notes: t-values in parentheses. The standard errors of the coefficients that have not been directly estimated are calculated with the formula of Klein (1953, p. 258). α_0 is set to 10.8, because this value gives the highest likelihood value of the AIDS Model.

Table 8.A8: Price elasticities of (final) profit function (restricted)

	P_c	P_a	P_v	P_L
X_c	0.429	0.503	-0.567	-0.364
X_a	0.320	0.533	-0.735	-0.118
X_v	0.356	0.726	-1.081	-0.001
X_L	0.340	0.172	-0.002	-0.511

Table 8.A9: Price and income elasticities - AIDS model

	Price elasticities						Income
	Hicksian elasticities			Marshallian elasticities			elasticities
	P_m	P_a	P_L	P_m	P_a	P_L	Y
C_m	-0.554	0.144	0.409	-0.667	0.119	0.148	0.400
C_a	0.648	-0.782	0.134	0.503	-0.815	-0.200	0.511
C_L	0.176	0.014	-0.190	-0.194	-0.070	-1.045	1.308

Kapitel 9

How to Estimate the “Almost Ideal Demand System”

Arne Henningsen

veröffentlicht als Poster-Paper auf der 25. internationalen Konferenz der „International Association of Agricultural Economists“ (IAAE) in Durban, Südafrika, 16.–22.08.2003.

Abstract

The Almost Ideal Demand System (AIDS) proposed by Deaton and Muellbauer (1980a) is widely used in applied demand analysis, because it unifies almost all desirable properties. To avoid complicated non-linear estimation of the original AIDS most empirical studies use a linear approximation of AIDS (LA-AIDS) by approximating the translog price index with the Stone index.

However, this leads to a poor approximation of the original AIDS parameters. Moreover, econometric estimation is plagued by several additional problems and the LA-AIDS is no more an integrable demand system. Several scholars tried to circumvent these problems by deriving special formulas for elasticities or using different price indices. However, since none of these approaches solves all estimation problems and the LA-AIDS remains a non-integrable demand system, other authors suggested iterative linear estimation procedures estimating the original AIDS.

In this framework the poster presents the results of a Monte Carlo simulation estimating various AIDS and LA-AIDS approaches. These analyses imply following results: (a) The LA-AIDS approaches using the Tornqvist and Laspeyres price indices lead generally to good local approximations. (b) The non-linear AIDS can be very accurately estimated by the iterative linear estimation procedure suggested by Michalek and Keyzer (1992). (c) If a non-linear AIDS is estimated, it is very important to choose the value for the α_0 that gives the best fit to the model. (d) Mean-scaling of prices does not bias the results and therefore this can be used to dramatically simplify the calculation of elasticities.

Keywords: Demand Analysis, AIDS, Estimation

9.1 Introduction

The Almost Ideal Demand System (AIDS) proposed by Deaton and Muellbauer (1980a,b) is widely used in applied demand analysis. As Deaton and Muellbauer stated it unifies almost all theoretically and empirically desirable properties: (a) It is an arbitrary first-order approximation to any demand system; (b) it satisfies the axioms of choice; (c) it aggregates exactly over consumers while still allowing nonlinear Engel curves, (d) the homogeneity and symmetry property can be tested and imposed by simple parameter restrictions and (e) the demand equations become linear if the Translog price index is approximated e.g. by the Stone index (see Deaton and Muellbauer, 1980a, p. 312). However, in empirical demand analysis especially the last property leads to problems.

In these premises the poster aims at persons that are interested in applied demand analysis. It should provide a broad overview over problems of the Almost Ideal Demand System (AIDS) and its linear approximation (LA-AIDS) and, most important, it gives practical hints to circumvent these problems.

9.2 Problems and Possible Solutions

Green and Alston (1990) were the first who showed that the suggested linearization of the AIDS leads in general to a poor approximation of the original AIDS-parameters. Thus, giving the fact that non-linear econometric estimation is often problematic it seems that Deaton and Muellbauer (1980a) were a little bit too optimistic at least regarding the last property. As a matter of fact since the pathbreaking paper of Deaton and Muellbauer (1980a) a bunch of literature evolved dealing with the problems of the econometric estimation of AIDS models. First, Green and Alston (1990) pointed out that deriving elasticities from a linear approximate AIDS (LA-AIDS) by differentiating the linearized share equations leads c.p. to a better approximation of the true AIDS-elasticities than deriving these directly from the non-linear share equations. However, for their better approximation they have to pay the price that their suggested LA-AIDS is no more an integrable demand system. Moreover, econometric estimation is plagued by several problems, e.g. simultaneity bias, errors-in-variables problem (Buse, 1994, p. 782) and units of measurement problem (Moschini, 1995). Second, Moschini (1995) suggested a LA-AIDS using the Paasche, Laspeyres and Tornqvist instead of the Stone price index substituting the Translog price index. While Moschini's approach solves some of the problems inherent in the LA-AIDS with the Stone price index, it remains a non-integrable system that is theoretically inconsistent. Third, Deaton and Muellbauer (1980a), Michalek and Keyzer (1992) and Pashardes (1993) propose different iterative linear estimation procedures to estimate the original non-linear AIDS including the Translog price index.

9.3 Monte Carlo Analysis

In this framework we performed a Monte Carlo simulation to evaluate the various AIDS and LA-AIDS approaches mentioned above: First, an AIDS for 4 goods was specified with "true" parameters that are in the range of typical results of empirical demand analyses. Second, exogenous variables were generated by (pseudo) random numbers and the "true" endogenous variables were calculated. Third, random errors were added to get "obtained" endogenous variables and the various AIDS and LA-AIDS approaches were applied to estimated coefficients, elasticities and welfare measures. The third step was repeated 1000 times. To enhance the generality of the results the Monte Carlo simulation was repeated under 48 different designs.

While our simulation is in several features similar to those of Buse (1993, 1994) and Buse and Chan (2000), it extends them in several aspects considerably: (a) simulated disturbances are more realistically generated according to the Dirichlet distribution, (b) additional designs (e.g. exogeneity of total expenditure) that are more realistic for several empirical demand analyses are applied, (c) additionally to coefficients and elasticities the accuracy of estimated welfare measures are examined, (d) the troublesome estimation of coefficient α_0 is investigated, (e) the iterative linear estimation procedures are systematically analyzed, and (f) the effects of mean-scaling of prices and expenditure are investigated.

9.4 Results

A selection of the results of our Monte Carlo analysis is presented in Tables 9.1 and 9.2. The mean square errors (MSE) of the estimated coefficients, elasticities and welfare measures were calculated for the different estimation methods and elasticity formulas. To summarize the results the (geometric) averages of the MSEs across the 48 different Monte Carlo designs were calculated. For simplification tables present "relative average MSEs" defined as the average MSEs divided by the average MSE of the full information maximum likelihood (FIML) estimation with the true α_0.

Table 9.1 shows our results of various non-linear estimation approaches. When the "true" α_0 is chosen a priori, the iterative method of Michalek and Keyzer (1992) is almost as good as the FIML or the non-linear SUR (Generalized Least Squares/Minimum Distance) estimation, while the iterative method of Deaton and Muellbauer (1980a) is slightly worse and the iterative method of Pashardes (1993) is considerably worse. In opposite to the first three estimation methods, which are nearly invariant to scaling of prices or expenditure, the estimation results of the latter two approaches can be improved by mean-scaling of prices.

If a wrong α_0 is chosen a priori (e.g. $\alpha_0 = 10$ instead of $\alpha_0 = 1$), the inaccuracy of the estimated coefficients, elasticities and welfare measures increases dramatically. Since in empirical demand analysis it is impossible to know the "true" α_0 a priori, we analyzed a procedure proposed by Michalek and Keyzer (1992). They proposed to estimate the AIDS with different values for α_0 and take the value that gives the best fit to the model. Our results show that although the estimated coefficients are quite inaccurate this procedure is able to accurately estimate elasticities and welfare measures.

Table 9.1: Results for non-linear AIDS estimations (selection)

Estimation Method[1]	α_0	Mean-scaling[2]	Valid Regr. [%]	Relative average MSE of elasticities	coefficients	welf. meas.
FIML	true[3]	–,P,P+X	99.95	1.00	1.00	1.00
nSUR	true[3]	–,P,P+X	100.00	0.99	1.00	1.00
M+K	true[3]	–,P,P+X	100.00	1.05	1.06	1.06
D+M	true[3]	–	100.00	1.29	1.41	1.23
D+M	true[3]	P,P+X	100.00	1.19	1.26	1.14
Pash.	true[3]	–	100.00	3.04	4.09	4.25
Pash.	true[3]	P,P+X	100.00	1.56	1.77	1.58
FIML	10	–	99.90	2.57	2104.84	2.39
nSUR	10	–	100.00	2.26	2144.36	2.11
M+K	10	–	99.88	4.18	2115.33	4.19
D+M	10	–	100.00	1.87	2718.51	2.56
FIML	best fit	–	99.86	1.01	442.72	1.03
nSUR	best fit	–	99.94	1.01	553.98	1.03
M+K	best fit	–	100.00	1.04	200.25	1.06
D+M	best fit	–	100.00	1.07	208.04	1.09

Source: Own Calculations

[1] FIML = Full Information Maximum Likelihood; nSUR = non-linear SUR (Generalized Least Squares, Minimum Distance); M+K = Iterative SUR estimation of Michalek and Keyzer (1992); D+M = Iterative SUR estimation of Deaton and Muellbauer (1980a); Pash. = Iterative SUR estimation of Pashardes (1993).

[2] – = no mean-scaling; P = mean-scaling of prices; P+X = mean-scaling of prices and total expenditure

[3] in the original model the 'true' α_0 is 1. If prices or total expenditure are mean-scaled, it is adjusted to take into account this change(s) of units of measurement.

Table 9.2 shows our results for various LA-AIDS approaches. Four different price indices, mean scaling of prices and both prices and expenditure, two different estimation methods and several different formulas for calculating elasticities were analyzed. Our results show that the LA-AIDS is generally not a good approximation of the AIDS. However, if both prices and expenditure are mean-scaled, the accuracy of the results improves, especially if the Tornqvist index is used. The effect of the estimation method is ambiguous and mainly does not influence the results heavily. If the exogenous variables are not mean-scaled, the

Table 9.2: Results for LA-AIDS estimations (selection)

Price index[1]	Mean-scaling[2]	Estim. Method	Elast. formula[3]	Mean relative MSE of elasticities	coefficients	welf. meas.
S	–	SUR	AIDS	23.41	40.31	17.06
S	–	3SLS	AIDS	26.38	44.48	19.69
S,P	P[4]	SUR	AIDS	13.83	16.56	176.49
S,P	P[4]	3SLS	AIDS	14.71	17.64	185.72
S,P	P+X	SUR	AIDS	1.28	38.68	1.43
S,P	P+X	3SLS	AIDS	1.19	38.71	1.48
S	–	SUR	G&A	1.41	see AIDS	
S	–	3SLS	G&A	1.24	see AIDS	
S	–	SUR	Goddard	1.59	see AIDS	
S	–	3SLS	Goddard	1.85	see AIDS	
S,P	P,P+X[4]	SUR	Godd., G&A	1.26	see AIDS	
S,P	P,P+X[4]	3SLS	Godd., G&A	1.16	see AIDS	
L	–	SUR,3SLS	AIDS	19.84	36.31	13.81
L	P	SUR,3SLS	AIDS	13.90	15.25	3.86
L	P+X	SUR,3SLS	AIDS	1.07	37.68	1.23
L	–,P,P+X	SUR,3SLS	Godd., G&A	1.07	see AIDS	
T	–,P	SUR	AIDS	14.20	17.19	178.53
T	–,P	3SLS	AIDS	14.36	17.37	180.07
T	P+X	SUR	AIDS	1.06	38.17	1.06
T	P+X	3SLS	AIDS	1.06	38.18	1.07
T	–,P,P+X	SUR	Godd., G&A	1.04	see AIDS	
T	–,P,P+X	3SLS	Godd., G&A	1.04	see AIDS	

Source: Own Calculations

[1] S = Stone, P = Paasche L = Laspeyres, T = Tornqvist

[2] – = no mean-scaling; P = mean-scaling of prices; P+X = mean-scaling of prices and total expenditure

[3] AIDS = original AIDS formula, G&A = formula of Green and Alston (1990), Godd. or Goddard = formula of Goddard (1983)

[4] including Paasche index with no mean-scaling

alternative elasticity formulas lead to much better results than the original AIDS formulas, but if both prices and expenditure are mean-scaled, the original AIDS formulas are almost as accurate as the alternative formulas.

9.5 Conclusions

Summarizing our results we conclude that acceptable estimates of elasticities can be obtained by a simple LA-AIDS estimation with Tornqvist or Laspeyres price index. If prices and expenditure are mean-scaled, the results of the LA-AIDS with Tornqvist index might even be used for welfare analysis.

However, the LA-AIDS remains only a theoretically inconsistent (local) approximation of the AIDS. Therefore, the further efforts of estimating the full AIDS might be worth it — in particular because the non-linear estimation can be accurately done by the iterative linear estimation procedure suggested by Michalek and Keyzer (1992). If a full AIDS is estimated, it is very important to choose the value for the α_0 that gives the best fit to the model.

Furthermore, in contrast to Buse and Chan (2000) our results show that mean-scaling does not bias the results. Therefore, mean-scaling of prices can be used to simplify the calculation of elasticities (at mean values) dramatically, because in this case several parts of the original AIDS formula become zero and analogically the complex formulas of Green and Alston (1991) and Buse (1994) reduce to the simple formula suggested by Goddard (1983).

Bibliography

Buse, A. (1993). Evaluating the Linearized Almost Ideal Demand System. Research Paper No. 93-18, Department of Economics, University of Alberta, Edmonton.

Buse, A. (1994). Evaluating the Linearized Almost Ideal Demand System. *American Journal of Agricultural Economics*, 76:781–793.

Buse, A. and Chan, W. H. (2000). Invariance, price indices and estimation in Almost Ideal Demand Systems. *Empirical Economics*, 25:519–539.

Deaton, A. and Muellbauer, J. (1980a). An Almost Ideal Demand System. *The American Economic Review*, 70:312–326.

Deaton, A. and Muellbauer, J. (1980b). *Economics and Consumer Behaviour*. Cambridge University Press, Cambridge.

Goddard, D. (1983). An analysis of Canadian aggregate demand for food at home and away from home. *Canadian Journal of Agricultural Economics*, 31:289–318.

Green, R. and Alston, J. M. (1990). Elasticities in AIDS models. *American Journal of Agricultural Economics*, 72:442–445.

Green, R. and Alston, J. M. (1991). Elasticities in AIDS models: A clarification and extension. *American Journal of Agricultural Economics*, 73:874–875.

Michalek, J. and Keyzer, M. A. (1992). Estimation of a two-stage LES-AIDS consumer demand system for eight EC countries. *European Review of Agricultural Economics*, 19(2):137–163.

Moschini, G. (1995). Units of measurement and the Stone index in demand system estimation. *American Journal of Agricultural Economics*, 77:63–68.

Pashardes, P. (1993). Bias in estimating the Almost Ideal Demand System with the Stone index approximation. *The Economic Journal*, 103:908–915.

Kapitel 10

systemfit: A Package to Estimate Simultaneous Equation Systems in R

Arne Henningsen und Jeff D. Hamann

zur Veröffentlichung akzeptiert mit Änderungsauflagen in: Journal of Statistical Software

Abstract

Many statistical analyses are based on models containing systems of structurally related equations. In cases where cross-equation disturbances are correlated, full information methods are required (Zellner, 1962). If exogenous variables are stochastically dependent on the disturbances in the system, then instrumental variable estimation methods should be used (Zellner and Theil, 1962) The package **systemfit** provides the capability to estimate systems of linear equations within the R programming environment.

Keywords: R, simultaneous equations systems, seemingly unrelated regression, two-stage least squares, three-stage least squares

10.1 Introduction

Many theoretical models that are econometrically estimated consist of more than one equation. The disturbance terms of these equations are likely to be contemporaneously correlated, because unconsidered factors that influence the disturbance term in one equation probably influence the disturbance terms in other equations. Ignoring this contemporaneous correlation and estimating these equations separately leads to inefficient parameter estimates. However, estimating all equations simultaneously, taking the covariance structure of the residuals into account, leads to efficient estimates. This estimation procedure is generally called "Seemingly Unrelated Regression" (SUR) (Zellner, 1962). Another reason to estimate an equation system simultaneously are cross-equation parameter restrictions.[1] These restrictions can be tested and/or imposed only in a simultaneous estimation approach.

Furthermore, these models can contain variables that appear on the left-hand side in one equation and on the right-hand side of another equation. Ignoring the endogeneity of these variables can lead to inconsistent parameter estimates. This simultaneity bias can be corrected for in each equation by applying a "Two-Stage Least Squares" (2SLS) method or for all equations simultaneously when combined with SUR resulting in a "Three-Stage Least Squares" (3SLS) estimation of the system of equations.

The **systemfit** package provides the capability to estimate linear equation systems in R (R Development Core Team, 2005). Although linear equation systems can be estimated with several other statistical and econometric software packages (e.g. SAS, EViews, TSP), **systemfit** has several advantages. First, all estimation procedures are publicly available in the source code. Second, the estimation algorithms can be easily modified to meet specific requirements.

[1]Especially the economic theory suggests many cross-equation parameter restrictions (e.g. the symmetry restriction in demand models).

Third, the (advanced) user can control estimation details generally not available in other software packages by overriding reasonable defaults.

In Section 10.2 we introduce the statistical background of estimating equation systems. The implementation of the statistical procedures in R is shortly explained in Section 10.3. Section 10.4 demonstrates how to run **systemfit** and how some of the features presented in the previous section can be utilized. In Section 10.5 the reliability of the results from **systemfit** are presented. Finally, a summary and outlook are presented in Section 10.6.

10.2 Statistical background

In this section we provide the statistical background of the functionality provided by the **systemfit** package. After introducing notations and assumptions, we provide the formulas to estimate systems of linear equations. We then demonstrate how to impose linear restrictions on parameters. Finally, we present additional relevant issues about estimation of equation systems.

Consider a system of G equations, where the ith equation is of the form

$$y_i = X_i \beta_i + u_i, \quad i = 1, 2, \ldots, G \tag{10.1}$$

where y_i is a vector of the dependent variable, X_i is a matrix of the exogenous variables, β_i is the coefficient vector and u_i is a vector of the disturbance terms of the ith equation.

We can write the "stacked" system as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_G \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_G \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_G \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_G \end{bmatrix} \tag{10.2}$$

or more simply as

$$y = X\beta + u \tag{10.3}$$

We assume that there is no correlation of the disturbance terms across observations:

$$E\left(u_{it}\, u_{jt^*}\right) = 0 \;\forall\; t \neq t^* \tag{10.4}$$

where i and j indicate the equation number and t and t^* denote the observation number.

However, we explicitly allow for contemporaneous correlation:

$$E\left(u_{it}\, u_{jt}\right) = \sigma_{ij} \tag{10.5}$$

Thus, the covariance matrix of the total system is

$$E\left(u\, u'\right) = \Omega = \Sigma \otimes I \tag{10.6}$$

where $\Sigma = [\sigma_{ij}]$ is the residual covariance matrix and I is an identity matrix.

10.2.1 Estimation

10.2.1.1 Ordinary least squares (OLS)

The Ordinary Least Squares (OLS) estimator of the system is obtained by

$$\widehat{\beta}_{OLS} = \left(X'X\right)^{-1} X'y \tag{10.7}$$

These estimates are efficient only if the disturbance terms are not contemporaneously correlated, which means $\sigma_{ij} = 0 \;\forall\; i \neq j$. If the whole system is treated as one single equation, the covariance matrix of the estimated parameters is

$$\operatorname{Cov}\left[\widehat{\beta}_{OLS}\right] = \sigma^2 \left(X'X\right)^{-1} \tag{10.8}$$

with $\sigma^2 = E\left(u'u\right)$. This assumes that the disturbances of all equations have the same variance.

If the disturbance terms of the individual equations are allowed to have different variances, the covariance matrix of the estimated parameters is

$$\operatorname{Cov}\left[\widehat{\beta}_{OLS}\right] = \left(X'\Omega^{-1}X\right)^{-1} \tag{10.9}$$

with $\Omega = \Sigma \otimes I$, $\sigma_{ij} = 0 \;\forall\; i \neq j$ and $\sigma_{ii} = E\left(u_i'u_i\right)$.

If no cross-equation parameter restrictions are imposed, the simultaneous OLS estimation of the system leads to the same parameter estimates as an equation-wise OLS estimation. The covariance matrix of the parameters from an equation-wise OLS estimation is equal to the covariance matrix obtained by equation (10.9).

10.2.1.2 Weighted least squares (WLS)

The Weighted Least Squares (WLS) estimator of the system is obtained by

$$\widehat{\beta}_{WLS} = \left(X'\Omega^{-1}X \right)^{-1} X'\Omega^{-1}y \tag{10.10}$$

with $\Omega = \Sigma \otimes I$, $\sigma_{ij} = 0 \; \forall \; i \neq j$ and $\sigma_{ii} = E\left(u_i' u_i\right)$. Like the OLS estimates these estimates are only efficient if the disturbance terms are not contemporaneously correlated. The covariance matrix of the estimated parameters is

$$\text{Cov}\left[\widehat{\beta}_{WLS}\right] = \left(X'\Omega^{-1}X \right)^{-1} \tag{10.11}$$

If no cross-equation parameter restrictions are imposed, the parameter estimates are equal to the OLS estimates.

10.2.1.3 Seemingly unrelated regression (SUR)

When the disturbances are contemporaneously correlated, a Generalized Least Squares (GLS) estimation leads to efficient parameter estimates. In this case, the GLS is generally called "Seemingly Unrelated Regression" (SUR) (Zellner, 1962). It should be noted that while an unbiased OLS or WLS estimation requires only that the regressors and the disturbance terms of each single equation are uncorrelated ($E\left[u_i|X_i\right] = 0 \; \forall \; i$), a consistent SUR estimation requires that all disturbance terms and all regressors are uncorrelated ($E\left[u|X\right] = 0$).

The SUR estimator can be obtained by:

$$\widehat{\beta}_{SUR} = \left(X'\Omega^{-1}X \right)^{-1} X'\Omega^{-1}y \tag{10.12}$$

with $\Omega = \Sigma \otimes I$ and $\sigma_{ij} = E\left(u_i' u_j\right)$. And the covariance matrix of the estimated parameters is

$$\text{Cov}\left[\widehat{\beta}_{SUR}\right] = \left(X'\Omega^{-1}X \right)^{-1} \tag{10.13}$$

10.2.1.4 Two-stage least squares (2SLS)

If the regressors of one or more equations are correlated with the disturbances ($E\left(u_i|X_i\right) \neq 0$), the estimated coefficients are biased. This can be circumvented by an instrumental variable (IV) two-stage least squares (2SLS) estimation. The instrumental variables for each equation H_i can be either different or identical for all equations. The instrumental variables of each equation may not be correlated with the disturbance terms of the corresponding equation ($E(u_i|H_i) = 0$).

At the first stage new ('fitted') regressors are obtained by

$$\widehat{X_i} = H_i \left(H_i' H_i\right)^{-1} H_i' X \tag{10.14}$$

At the second stage the unbiased two-stage least squares estimates of β are obtained by:

$$\widehat{\beta}_{2SLS} = \left(\widehat{X}'\widehat{X}\right)^{-1} \widehat{X}'y \tag{10.15}$$

If the whole system is treated as one single equation, the covariance matrix of the estimated parameters is

$$\text{Cov}\left[\widehat{\beta}_{2SLS}\right] = \sigma^2 \left(\widehat{X}'\widehat{X}\right)^{-1} \tag{10.16}$$

with $\sigma^2 = E(u'u)$. If the disturbance terms of the individual equations are allowed to have different variances, the covariance matrix of the estimated parameters is

$$\text{Cov}\left[\widehat{\beta}_{2SLS}\right] = \left(\widehat{X}'\Omega^{-1}\widehat{X}\right)^{-1} \tag{10.17}$$

with $\Omega = \Sigma \otimes I$, $\sigma_{ij} = 0 \;\forall\; i \neq j$ and $\sigma_{ii} = E(u_i' u_i)$.

10.2.1.5 Weighted two-stage least squares (W2SLS)

The Weighted Two-Stage Least Squares (W2SLS) estimator of the system is obtained by

$$\widehat{\beta}_{W2SLS} = \left(\widehat{X}'\Omega^{-1}\widehat{X}\right)^{-1} \widehat{X}'\Omega^{-1}y \tag{10.18}$$

with $\Omega = \Sigma \otimes I$, $\sigma_{ij} = 0 \;\forall\; i \neq j$ and $\sigma_{ii} = E(u_i' u_i)$. The covariance matrix of the estimated parameters is

$$\text{Cov}\left[\widehat{\beta}_{W2SLS}\right] = \left(\widehat{X}'\Omega^{-1}\widehat{X}\right)^{-1} \tag{10.19}$$

10.2.1.6 Three-stage least squares (3SLS)

If the regressors are correlated with the disturbances ($E(u|X) \neq 0$) and the disturbances are contemporaneously correlated, a Generalized Least Squares (GLS) version of the two-stage least squares estimation leads to consistent and efficient estimates. This estimation procedure is generally called "Three-stage Least Squares" (3SLS) (Zellner and Theil, 1962).

The standard 3SLS estimator can be obtained by:

$$\widehat{\beta}_{3SLS} = \left(\widehat{X}'\Omega^{-1}\widehat{X}\right)^{-1} \widehat{X}'\Omega^{-1}y \tag{10.20}$$

with $\Omega = \Sigma \otimes I$ and $\sigma_{ij} = E\left(u_i' u_j\right)$. Its covariance matrix is:

$$\text{Cov}\left[\widehat{\beta}_{3SLS}\right] = \left(\widehat{X}'\Omega^{-1}\widehat{X}\right)^{-1} \tag{10.21}$$

While an unbiased 2SLS or W2SLS estimation requires only that the instrumental variables and the disturbance terms of each single equation are uncorrelated ($E\left[u_i|H_i\right]) = 0 \;\forall\; i$), Schmidt (1990) points out that this estimator is only consistent if all disturbance terms and all instrumental variables are uncorrelated ($E\left[u|H\right]) = 0$) with

$$H = \begin{bmatrix} H_1 & 0 & \cdots & 0 \\ 0 & H_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & H_G \end{bmatrix} \tag{10.22}$$

Since there might be occasions where this cannot be avoided, Schmidt (1990) analyses other approaches to obtain 3SLS estimators:

One of these approaches is based on instrumental variable estimation (3SLS-IV):

$$\widehat{\beta}_{3SLS-IV} = \left(\widehat{X}'\Omega^{-1}X\right)^{-1} \widehat{X}'\Omega^{-1}y \tag{10.23}$$

The covariance matrix of this 3SLS-IV estimator is:

$$\text{Cov}\left[\widehat{\beta}_{3SLS-IV}\right] = \left(\widehat{X}'\Omega^{-1}X\right)^{-1} \tag{10.24}$$

Another approach is based on the Generalized Method of Moments (GMM) estimator (3SLS-GMM):

$$\widehat{\beta}_{3SLS-GMM} = \left(X'H\left(H'\Omega H\right)^{-1}H'X\right)^{-1} X'H\left(H'\Omega H\right)^{-1}H'y \tag{10.25}$$

The covariance matrix of the 3SLS-GMM estimator is:

$$\text{Cov}\left[\widehat{\beta}_{3SLS-GMM}\right] = \left(X'H\left(H'\Omega H\right)^{-1}H'X\right)^{-1} \tag{10.26}$$

A fourth approach developed by Schmidt (1990) himself is:

$$\widehat{\beta}_{3SLS-Schmidt} = \left(\widehat{X}'\Omega^{-1}\widehat{X}\right)^{-1} \widehat{X}'\Omega^{-1}H\left(H'H\right)^{-1}H'y \tag{10.27}$$

The covariance matrix of this estimator is:

$$\text{Cov}\left[\widehat{\beta}_{3SLS-Schmidt}\right] = \left(\widehat{X}'\Omega^{-1}\widehat{X}\right)^{-1} \widehat{X}'\Omega^{-1}H\left(H'H\right)^{-1}H'\Omega H\left(H'H\right)^{-1}H'\Omega^{-1}\widehat{X}\left(\widehat{X}'\Omega^{-1}\widehat{X}\right)^{-1} \tag{10.28}$$

The econometrics software EViews uses following approach:

$$\widehat{\beta}_{3SLS-EViews} = \widehat{\beta}_{2SLS} + \left(\widehat{X}'\Omega^{-1}\widehat{X}\right)^{-1} \widehat{X}'\Omega^{-1}\left(y - X\widehat{\beta}_{2SLS}\right) \tag{10.29}$$

where $\widehat{\beta}_{2SLS}$ is the two-stage least squares estimator as defined by (10.15). EViews uses the standard 3SLS formula (10.21) to calculate the covariance matrix of the 3SLS estimator.

If the same instrumental variables are used in all equations ($H_1 = H_2 = \ldots = H_G$), all the above mentioned approaches lead to identical parameter estimates. However, if this is not the case, the results depend on the method used (Schmidt, 1990). The only reason to use different instruments for different equations is a correlation of the instruments of one equation with the disturbance terms of another equation. Otherwise, one could simply use all instruments in every equation (Schmidt, 1990). In this case, only the 3SLS-GMM (10.25) and the 3SLS estimator developed by Schmidt (1990) (10.27) are consistent.

10.2.2 Imposing linear restrictions

It is common to perform hypothesis tests by imposing restrictions on the parameter estimates. There are two ways to impose linear parameter restrictions. First, a matrix T can be specified that

$$\beta = T \cdot \beta^* \tag{10.30}$$

where β^* is a vector of restricted (linear independent) coefficients, and T is a matrix with the number of rows equal to the number of unrestricted coefficients (β) and the number of columns equal to the number of restricted coefficients (β^*). T can be used to map each unrestricted coefficient to one or more restricted coefficients.

To impose these restrictions, the X matrix is (post-)multiplied by this T matrix.

$$X^* = X \cdot T \tag{10.31}$$

Then, X^* is substituted for X and a standard estimation as described in the previous section is done (equations 10.7–10.29). This results in the linear independent parameter estimates β^* and their covariance matrix. The original parameters can be obtained by equation (10.30) and the covariance matrix of the original parameters can be obtained by:

$$\text{Cov}\left[\widehat{\beta}\right] = T \cdot \text{Cov}\left[\widehat{\beta}^*\right] \cdot T' \tag{10.32}$$

The second way to impose linear parameter restrictions can be formulated by

$$R\beta^0 = q \tag{10.33}$$

where β^0 is the vector of the restricted coefficients, and R and q are a matrix and vector, respectively, to impose the restrictions (see Greene, 2003, p. 100). Each linear independent restriction is represented by one row of R and the corresponding element of q.

The first way is less flexible than this latter one[2], but the first way is preferable if equality constraints for coefficients across many equations of the system are imposed. Of course, these restrictions can be also imposed using the latter method. However, while the latter method increases the dimension of the matrices to be inverted during estimation, the first reduces it. Thus, in some cases the latter way leads to estimation problems (e.g. (near) singularity of the matrices to be inverted), while the first does not.

These two methods can be combined. In this case the restrictions imposed using the latter method are imposed on the linear independent parameters due to the restrictions imposed using the first method:

$$R\beta^{*0} = q \tag{10.34}$$

where β^{*0} is the vector of the restricted β^* coefficients.

10.2.2.1 Restricted OLS estimation

The OLS estimator restricted by $R\beta^0 = q$ can be obtained by

$$\begin{bmatrix} \widehat{\beta}^0_{OLS} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} X'X & R' \\ R & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} X'y \\ q \end{bmatrix} \tag{10.35}$$

[2]While restrictions like $\beta_1 = 2\beta_2$ can be imposed by both methods, restrictions like $\beta_1 + \beta_2 = 4$ can be imposed only by the second method.

where λ is a vector of the Lagrangean multipliers of the restrictions. If the whole system is treated as one single equation, the covariance matrix of the estimated parameters is

$$\text{Cov}\begin{bmatrix} \widehat{\beta}^0_{OLS} \\ \widehat{\lambda} \end{bmatrix} = \sigma^2 \begin{bmatrix} X'X & R' \\ R & 0 \end{bmatrix}^{-1} \tag{10.36}$$

with $\sigma^2 = E\left(u'u\right)$. If the disturbance terms of the individual equations are allowed to have different variances, the covariance matrix of the estimated parameters is

$$\text{Cov}\begin{bmatrix} \widehat{\beta}^0_{OLS} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} X'\Omega^{-1}X & R' \\ R & 0 \end{bmatrix}^{-1} \tag{10.37}$$

with $\Omega = \Sigma \otimes I$, $\sigma_{ij} = 0 \; \forall \; i \neq j$ and $\sigma_{ii} = E\left(u_i'u_i\right)$.

10.2.2.2 Restricted WLS estimation

The WLS estimator restricted by $R\beta^0 = q$ can be obtained by

$$\begin{bmatrix} \widehat{\beta}^0_{WLS} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} X'\Omega^{-1}X & R' \\ R & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} X'\Omega^{-1}y \\ q \end{bmatrix} \tag{10.38}$$

with $\Omega = \Sigma \otimes I$, $\sigma_{ij} = 0 \; \forall \; i \neq j$ and $\sigma_{ii} = E\left(u_i'u_i\right)$. The covariance matrix of the estimated parameters is

$$\text{Cov}\begin{bmatrix} \widehat{\beta}^0_{WLS} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} X'\Omega^{-1}X & R' \\ R & 0 \end{bmatrix}^{-1} \tag{10.39}$$

10.2.2.3 Restricted SUR estimation

The SUR estimator restricted by $R\beta^0 = q$ can be obtained by

$$\begin{bmatrix} \widehat{\beta}^0_{SUR} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} X'\Omega^{-1}X & R' \\ R & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} X'\Omega^{-1}y \\ q \end{bmatrix} \tag{10.40}$$

with $\Omega = \Sigma \otimes I$ and $\sigma_{ij} = E\left(u_i'u_j\right)$. The covariance matrix of the estimated parameters is

$$\text{Cov}\begin{bmatrix} \widehat{\beta}^0_{SUR} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} X'\Omega^{-1}X & R' \\ R & 0 \end{bmatrix}^{-1} \tag{10.41}$$

10.2.2.4 Restricted 2SLS estimation

The 2SLS estimator restricted by $R\beta^0 = q$ can be obtained by

$$\begin{bmatrix} \widehat{\beta}^0_{2SLS} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{X}'\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \widehat{X}'y \\ q \end{bmatrix} \tag{10.42}$$

If the whole system is treated as one single equation, the covariance matrix of the estimated parameters is

$$\text{Cov} \begin{bmatrix} \widehat{\beta}^0_{2SLS} \\ \widehat{\lambda} \end{bmatrix} = \sigma^2 \begin{bmatrix} \widehat{X}'\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \tag{10.43}$$

with $\sigma^2 = E\left(u'u\right)$. If the disturbance terms of the individual equations are allowed to have different variances, the covariance matrix of the estimated parameters is

$$\text{Cov} \begin{bmatrix} \widehat{\beta}^0_{2SLS} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{X}'\Omega^{-1}\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \tag{10.44}$$

with $\Omega = \Sigma \otimes I$, $\sigma_{ij} = 0 \;\forall\; i \neq j$ and $\sigma_{ii} = E\left(u_i'u_i\right)$.

10.2.2.5 Restricted W2SLS estimation

The W2SLS estimator restricted by $R\beta^0 = q$ can be obtained by

$$\begin{bmatrix} \widehat{\beta}^0_{W2SLS} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{X}'\Omega^{-1}\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \widehat{X}'\Omega^{-1}y \\ q \end{bmatrix} \tag{10.45}$$

with $\Omega = \Sigma \otimes I$, $\sigma_{ij} = 0 \;\forall\; i \neq j$ and $\sigma_{ii} = E\left(u_i'u_i\right)$. The covariance matrix of the estimated parameters is

$$\text{Cov} \begin{bmatrix} \widehat{\beta}^0_{W2SLS} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{X}'\Omega^{-1}\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \tag{10.46}$$

10.2.2.6 Restricted 3SLS estimation

The standard 3SLS estimator restricted by $R\beta^0 = q$ can be obtained by

$$\begin{bmatrix} \widehat{\beta}^0_{3SLS} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{X}'\Omega^{-1}\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \widehat{X}'\Omega^{-1}y \\ q \end{bmatrix} \tag{10.47}$$

with $\Omega = \Sigma \otimes I$ and $\sigma_{ij} = E\left(u_i' u_j\right)$. The covariance matrix of this estimator is

$$\text{Cov}\begin{bmatrix} \widehat{\beta}^0_{3SLS} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{X}'\Omega^{-1}\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \tag{10.48}$$

The 3SLS-IV estimator restricted by $R\beta^0 = q$ can be obtained by

$$\begin{bmatrix} \widehat{\beta}^0_{3SLS-IV} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{X}'\Omega^{-1}X & R' \\ R & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \widehat{X}'\Omega^{-1}y \\ q \end{bmatrix} \tag{10.49}$$

with

$$\text{Cov}\begin{bmatrix} \widehat{\beta}^0_{3SLS-IV} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{X}'\Omega^{-1}\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \tag{10.50}$$

The restricted 3SLS-GMM estimator can be obtained by

$$\begin{bmatrix} \widehat{\beta}^0_{3SLS-GMM} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} X'H\left(H'\Omega H\right)^{-1}H'X & R' \\ R & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} X'H\left(H\Omega H\right)^{-1}H'y \\ q \end{bmatrix} \tag{10.51}$$

with

$$\text{Cov}\begin{bmatrix} \widehat{\beta}^0_{3SLS-GMM} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} X'H\left(H'\Omega H\right)^{-1}H'X & R' \\ R & 0 \end{bmatrix}^{-1} \tag{10.52}$$

The restricted 3SLS estimator based on the suggestion of Schmidt (1990) is:

$$\begin{bmatrix} \widehat{\beta}^0_{3SLS-Schmidt} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{X}'\Omega^{-1}\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \widehat{X}'\Omega^{-1}H\left(H'H\right)^{-1}H'y \\ q \end{bmatrix} \tag{10.53}$$

with

$$\begin{aligned} \text{Cov}\begin{bmatrix} \widehat{\beta}^0_{3SLS-Schmidt} \\ \widehat{\lambda} \end{bmatrix} &= \begin{bmatrix} \widehat{X}'\Omega^{-1}\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \\ &\cdot \begin{bmatrix} \widehat{X}'\Omega^{-1}H\left(H'H\right)^{-1}H'\Omega H\left(H'H\right)^{-1}H'\Omega^{-1}\widehat{X} & 0' \\ 0 & 0 \end{bmatrix}^{-1} \\ &\cdot \begin{bmatrix} \widehat{X}'\Omega^{-1}\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \end{aligned} \tag{10.54}$$

The econometrics software EViews calculates the restricted 3SLS estimator by:

$$\begin{bmatrix} \widehat{\beta}^0_{3SLS-EViews} \\ \widehat{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{X}'\Omega^{-1}\widehat{X} & R' \\ R & 0 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \widehat{X}'\Omega^{-1}\left(y - X\widehat{\beta}^0_{2SLS}\right) \\ q \end{bmatrix} \tag{10.55}$$

where $\widehat{\beta}^0_{2SLS}$ is the restricted 2SLS estimator calculated by equation (10.42). To calculate the covariance matrix EViews uses the standard formula of the restricted 3SLS estimator (10.48).

If the same instrumental variables are used in all equations ($H_1 = H_2 = \ldots = H_G$), all the above mentioned approaches lead to identical parameter estimates and identical covariance matrices of the estimated parameters.

10.2.3 Residual covariance matrix

Since the true residuals of the estimated equations are generally not known, the true covariance matrix of the residuals cannot be determined. Thus, this covariance matrix must be calculated from the *estimated* residuals. Generally, the estimated covariance matrix of the residuals ($\widehat{\Sigma} = [\widehat{\sigma}_{ij}]$) can be calculated from the residuals of a first-step OLS or 2SLS estimation. The following formula is often applied:

$$\widehat{\sigma}_{ij} = \frac{\widehat{u}_i'\widehat{u}_j}{T} \tag{10.56}$$

where T is the number of observations in each equation. However, in finite samples this estimator is biased, because it is not corrected for degrees of freedom. The usual single-equation procedure to correct for degrees of freedom cannot always be applied, because the number of regressors in each equation might differ. Two alternative approaches to calculate the residual covariance matrix are

$$\widehat{\sigma}_{ij} = \frac{\widehat{u}_i'\widehat{u}_j}{\sqrt{(T - K_i) \cdot (T - K_j)}} \tag{10.57}$$

and

$$\widehat{\sigma}_{ij} = \frac{\widehat{u}_i'\widehat{u}_j}{T - \max(K_i, K_j)} \tag{10.58}$$

where K_i and K_j are the number of regressors in equation i and j, respectively. However, these formulas yield unbiased estimators only if $K_i = K_j$ (Judge *et al.*, 1985, p. 469).

A further approach to obtain the estimated residual covariance matrix is (Zellner and Huang, 1962, p. 309)

$$\widehat{\sigma}_{ij} = \frac{\widehat{u}_i' \widehat{u}_j}{T - K_i - K_j + tr\left[X_i \left(X_i' X_i\right)^{-1} X_i' X_j \left(X_j' X_j\right)^{-1} X_j'\right]} \tag{10.59}$$

$$= \frac{\widehat{u}_i' \widehat{u}_j}{T - K_i - K_j + tr\left[\left(X_i' X_i\right)^{-1} X_i' X_j \left(X_j' X_j\right)^{-1} X_j' X_i\right]} \tag{10.60}$$

This yields an unbiased estimator for all elements of $\widehat{\Sigma}$, but even if $\widehat{\Sigma}$ is an unbiased estimator of Σ, its inverse $\widehat{\Sigma}^{-1}$ is not an unbiased estimator of Σ^{-1} (Theil, 1971, p. 322). Furthermore, the covariance matrix calculated by (10.59) is not necessarily positive semidefinite (Theil, 1971, p. 322). Hence, "it is doubtful whether [this formula] is really superior to [(10.56)]" (Theil, 1971, p. 322).

The WLS, SUR, W2SLS and 3SLS parameter estimates are consistent if the estimated residual covariance matrix is calculated using the residuals from a first-step OLS or 2SLS estimation. There exists also an alternative slightly different approach.[3] This alternative approach uses the residuals of a first-step OLS or 2SLS estimation to apply a WLS or W2SLS estimation on a second step. Then, it calculates the residual covariance matrix from the residuals of the second-step estimation to estimates the model by SUR or 3SLS in a third step. If no cross-equation restrictions are imposed, the parameter estimates of OLS and WLS as well as 2SLS and W2SLS are identical. Hence, in this case both approaches generate the same results.

It is also possible to iterate WLS, SUR, W2SLS and 3SLS estimations. At each iteration the residual covariance matrix is calculated from the residuals of the previous iteration. If equation (10.56) is applied to calculate the estimated residual covariance matrix, an iterated SUR estimation converges to maximum likelihood (Greene, 2003, p. 345).

In some uncommon cases, for instance in pooled estimations, where the coefficients are restricted to be equal in all equations, the means of the residuals of each equation are not equal to zero ($\overline{\widehat{u}}_i \neq 0$). Therefore, it might be argued that the residual covariance matrix should be calculated by subtracting the means from the residuals and substituting $\widehat{u}_i - \overline{\widehat{u}}_i$ for $\widehat{u}_i$ in (10.56–10.59).

[3]For instance, this approach is applied by the command "TSCS" of the software LIMDEP that carries out SUR estimations in which all coefficient vectors are constrained to be equal (Greene, 2006).

10.2.4 Degrees of freedom

To our knowledge the question about how to determine the degrees of freedom for single-parameter t tests is not comprehensively discussed in the literature. While sometimes the degrees of freedom of the entire system (total number of observations in all equations minus total number of estimated parameters) are applied, in other cases the degrees of freedom of each single equation (number of observations in the equations minus number of estimated parameters in the equation) are used. Asymptotically, this distinction does not make a difference. However, in many empirical applications, the number of observations of each equation is rather small, and therefore, it matters.

If a system of equations is estimated by an unrestricted OLS and the covariance matrix of the parameters is calculated by (10.9), the estimated parameters and their standard errors are identical to an equation-wise OLS estimation. In this case, it is reasonable to use the degrees of freedom of each single equation, because this yields the same p values as the equation-wise OLS estimation.

In contrast, if a system of equations is estimated with many cross-equation restrictions and the covariance matrix of an OLS estimation is calculated by (10.8), the system estimation is similar to a single equation estimation. Therefore, in this case, it seems to be reasonable to use the degrees of freedom of the entire system.

10.2.5 Goodness of fit

The goodness of fit of each single equation can be measured by the traditional R^2 values:

$$R_i^2 = 1 - \frac{\widehat{u}_i' \widehat{u}_i}{(y_i - \overline{y_i})'(y_i - \overline{y_i})} \tag{10.61}$$

where R_i^2 is the R^2 value of the ith equation and $\overline{y_i}$ is the mean value of y_i.

The goodness of fit of the whole system can be measured by the McElroy's R^2 value (McElroy, 1977):

$$R_*^2 = 1 - \frac{\widehat{u}' \widehat{\Omega}^{-1} \widehat{u}}{y' \left(\widehat{\Sigma}^{-1} \otimes \left(I - \frac{ii'}{T} \right) \right) y} \tag{10.62}$$

where T is the number of observations in each equation, I is an $T \times T$ identity matrix and i is a column vector of T ones.

10.2.6 Testing linear restrictions

Linear restrictions can be tested by an F test, Wald test or likelihood ratio (LR) test.

The F statistic for systems of equations is

$$F = \frac{(R\hat{\beta} - q)'(R(X'(\hat{\Sigma} \otimes I)^{-1}X)^{-1}R')^{-1}(R\hat{\beta} - q)/j}{\hat{u}'(\Sigma \otimes I)^{-1}\hat{u}/(M \cdot T - K)} \tag{10.63}$$

where j is the number of restrictions, M is the number of equations, T is the number of observations per equation, K is the total number of estimated coefficients, and $\hat{\Sigma}$ is the estimated residual covariance matrix used in the estimation. Under the null hypothesis, F has an F distribution with j and $M \cdot T - K$ degrees of freedom (Theil, 1971, p. 314).

The Wald statistic for systems of equations is

$$W = (R\hat{\beta} - q)'(R\widehat{Cov}[\hat{\beta}]R')^{-1}(R\hat{\beta} - q) \tag{10.64}$$

Asymptotically, W has a χ^2 distribution with j degrees of freedom under the null hypothesis (Greene, 2003, p. 347).

The LR statistic for systems of equations is

$$LR = T \cdot \left(log \left| \hat{\hat{\Sigma}}_r \right| - log \left| \hat{\hat{\Sigma}}_u \right| \right) \tag{10.65}$$

where T is the number of observations per equation, and $\hat{\hat{\Sigma}}_r$ and $\hat{\hat{\Sigma}}_u$ are the residual covariance matrices calculated by formula (10.56) of the restricted and unrestricted estimation, respectively. Asymptotically, LR has a χ^2 distribution with j degrees of freedom under the null hypothesis (Greene, 2003, p. 349).

10.2.7 Hausman test

Hausman (1978) developed a test for misspecification. The null hypothesis of the test is that all exogenous variables are uncorrelated with all disturbance terms. Under this hypothesis both the 2SLS and the 3SLS estimator are consistent but only the 3SLS estimator is (asymptotically) efficient. Under the alternative hypothesis the 2SLS estimator is consistent but the 3SLS estimator is inconsistent. The Hausman test statistic is,

$$m = \left(\hat{\beta}_{2SLS} - \hat{\beta}_{3SLS} \right)' \left(\text{Cov} \left[\hat{\beta}_{2SLS} \right] - \text{Cov} \left[\hat{\beta}_{3SLS} \right] \right) \left(\hat{\beta}_{2SLS} - \hat{\beta}_{3SLS} \right) \tag{10.66}$$

where $\hat{\beta}_{2SLS}$ and Cov $\left[\hat{\beta}_{2SLS}\right]$ are the estimated coefficient and covariance matrix from 2SLS estimation, and $\hat{\beta}_{3SLS}$ and Cov $\left[\hat{\beta}_{3SLS}\right]$ are the estimated coefficients and covariance matrix from 3SLS estimation. Under the null hypothesis this test statistic has a χ^2 distribution with degrees of freedom equal to the number of estimated parameters.

10.3 Source code

The **systemfit** package includes functions to estimate systems of equations (`systemfit`, `systemfitClassic`) and to test hypotheses in these systems (`ftest.systemfit`, `wald-test.systemfit`, `lrtest.systemfit`, `hausman.systemfit`). Furthermore, this package provides some helper functions e.g. to extract the estimated coefficients (`coef.systemfit`) or to calculate predicted values (`predict.systemfit`).

The source code of the **systemfit** is publicly available for download from "CRAN" (The Comprehensive R Archive Network, http://cran.r-project.org/src/contrib/Descriptions/systemfit.html). Since the whole package has more than 2,100 lines of code, it is not presented in this article. Furthermore, the code corresponds exactly to the procedures and formulas described in the previous section.

10.3.1 The basic function `systemfit`

The basic functionality of this package is provided by the function `systemfit`. This function estimates the equation system as described in sections 10.2.1. If parameter restrictions are provided, the formulas in section 10.2.2 are applied. Furthermore, the user can control several details of the estimation. For instance, the formula to calculate the residual covariance matrix (see section 10.2.3), the degrees of freedom for the t tests (see section 10.2.4), or the formula for the 3SLS estimation (see sections 10.2.1 and 10.2.2) can be specified by the user. The `systemfit` function returns many objects that users might be interest in. A complete list is available in the documentation of this function that is included in the package.

10.3.2 The wrapper function `systemfitClassic`

Furthermore, the **systemfit** package includes the function `systemfitClassic`. This is a wrapper function for `systemfit` that can be applied to (classical) panel-like data in long format if the regressors are the same for all equations. The data are reshaped and the formulas are modified to enable an estimation using the standard `systemfit` function. The user can specify whether the coefficients should be restricted to be equal in all equations.

10.3.3 Statistical tests

The statistical tests described in sections 10.2.6 and 10.2.7 are implemented as specified in these sections. The functions `ftest.systemfit`, `waldtest.systemfit` and `lrtest.systemfit` test linear restrictions on the estimated parameters. On the other hand, the function `hausman.systemfit` tests the consistency of the 3SLS estimator. All functions return the empirical test statistic, the degree(s) of freedom, and the p value.

10.3.4 Efficiency of the code

We have followed Bates (2004) to make the code faster and more stable. First, if a formula contains an inverse of a matrix that is post-multiplied by a vector, we use `solve(A,b)` instead of `solve(A) %*% b`. Second, we calculate crossproducts by `crossprod(X)` or `crossprod(X,y)` instead of `t(X) %*% X` or `t(X) %*% y`, respectively.

The matrix Ω^{-1} that is used to compute the estimated coefficients and their covariance matrix is of size $(G \cdot T) \times (G \cdot T)$ (see sections 10.2.1 and 10.2.2). In case of large data sets, this matrix Ω^{-1} gets really huge and needs a lot of memory. Therefore, we use the following transformation and compute $X'\Omega^{-1}$ by dividing the X matrix into submatrices, doing some calculations with these submatrices, adding up some of these submatrices, and finally putting the submatrices together:

$$X'\Omega^{-1} = \sum_{i=1} \begin{bmatrix} \sigma^{1i} X^1 \\ \sigma^{2i} X^2 \\ \vdots \\ \sigma^{Gi} X^G \end{bmatrix}' \tag{10.67}$$

where σ^{ij} are the elements of the matrix Σ^{-1}, and X^i is a submatrix of X that contains the rows that belong to the i's equation.

10.4 Using systemfit

In this section we demonstrate how to use the **systemfit** package. First, we show the standard usage of `systemfit` by a simple example. Second, several options that can be specified by the user are presented. Then, the wrapper function `systemfitClassic` is described. Finally, we demonstrate how to apply some statistical tests.

10.4.1 Standard usage of `systemfit`

As described in the previous section, equation systems can be econometrically estimated with the function `systemfit`. It is generally called by

```
> systemfit(method, eqns)
```

There are two mandatory arguments: `method` and `eqns`. The argument `method` is a string determining the estimation method. It must be either "OLS", "WLS", "SUR", "WSUR", "2SLS", "W2SLS", "3SLS", or "W3SLS". While six of these methods correspond to the estimation methods described in sections 10.2.1 and 10.2.2, the methods "WSUR" and "W3SLS" are "SUR" and "3SLS" estimations using the residual covariance matrices from "WLS" and "W2SLS" estimations, respectively (see section 10.2.3). The other mandatory argument, `eqns`, is a list of the equations to be estimated. Each equation is a standard formula in R and starts with a dependent variable on the left hand side. After a tilde ($\sim$) the regressors are listed, separated by plus signs[4].

The following demonstration uses an example taken from Kmenta (1986, p. 685). We want to estimate a small model of the US food market:

$$\texttt{consump} = \beta_1 + \beta_2 * \texttt{price} + \beta_3 * \texttt{income} \tag{10.68}$$

$$\texttt{consump} = \beta_4 + \beta_5 * \texttt{price} + \beta_6 * \texttt{farmPrice} + \beta_7 * \texttt{trend} \tag{10.69}$$

The first equation represents the demand side of the food market. Variable `consump` (food consumption per capita) is the dependant variable. The regressors are `price` (ratio of food prices to general consumer prices) and `income` (disposable income) as well as a constant. The second equation specifies the supply side of the food market. Variable `consump` is the dependant variable of this equation as well. The regressors are again `price` (ratio of food prices to general consumer prices) and a constant as well as `farmPrice` (ratio of preceding year's prices received by farmers) and `trend` (a time trend in years). These equations can be estimated as SUR in R by

```
> library(systemfit)
> data(Kmenta)
> attach(Kmenta)
> eqDemand <- consump ~ price + income
> eqSupply <- consump ~ price + farmPrice + trend
> fitsur <- systemfit("SUR", list(demand = eqDemand, supply = eqSupply))
```

[4]For details see the R help files to `formula`.

The first line loads the **systemfit** package. The second line loads example data that are included with the package. They are attached to the R search path in line three. In the fourth and fifth line, the demand and supply equations are specified, respectively[5]. Finally, in the last line, a seemingly unrelated regression is performed and the regression result is assigned to the variable `fitsur`.

Summary results can be printed by

```
> summary(fitsur)

systemfit results
method: SUR

        N DF      SSR     MSE    RMSE       R2   Adj R2
demand 20 17  65.6829 3.86370 1.96563 0.755019 0.726198
supply 20 16 104.0584 6.50365 2.55023 0.611888 0.539117

The covariance matrix of the residuals used for estimation
        demand  supply
demand 3.72539 4.13696
supply 4.13696 5.78444

The covariance matrix of the residuals
        demand  supply
demand 3.86370 4.92431
supply 4.92431 6.50365

The correlations of the residuals
         demand   supply
demand 1.000000 0.982348
supply 0.982348 1.000000

The determinant of the residual covariance matrix: 0.879285
OLS R-squared value of the system: 0.683453
McElroy's R-squared value for the system: 0.788722
```

[5] A regression constant is always implied if not explicitly omitted.

```
SUR estimates for 'demand' (equation 1)
Model Formula: consump ~ price + income

             Estimate Std. Error   t value Pr(>|t|)
(Intercept) 99.332894   7.514452 13.218913        0 ***
price       -0.275486   0.088509 -3.112513 0.006332  **
income        0.29855   0.041945  7.117605    2e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.96563 on 17 degrees of freedom
Number of observations: 20 Degrees of Freedom: 17
SSR: 65.682902 MSE: 3.8637 Root MSE: 1.96563
Multiple R-Squared: 0.755019 Adjusted R-Squared: 0.726198


SUR estimates for 'supply' (equation 2)
Model Formula: consump ~ price + farmPrice + trend

             Estimate Std. Error  t value Pr(>|t|)
(Intercept) 61.966166   11.08079 5.592215    4e-05 ***
price        0.146884   0.094435 1.555397 0.139408
farmPrice    0.214004   0.039868 5.367761  6.3e-05 ***
trend        0.339304   0.067911 4.996283 0.000132 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.550226 on 16 degrees of freedom
Number of observations: 20 Degrees of Freedom: 16
SSR: 104.05843 MSE: 6.503652 Root MSE: 2.550226
Multiple R-Squared: 0.611888 Adjusted R-Squared: 0.539117
```

First, the estimation method is reported and a few summary statistics for each equation are given. Then, some results regarding the whole equation system are printed: covariance matrix and correlations of the residuals, log of the determinant of the residual covariance matrix, R^2 value of the whole system, and McElroy's R^2 value. If the model is estimated by (W)SUR or (W)3SLS, the covariance matrix used for estimation is printed additionally. Finally, the estimation results of each equation are reported: formula of the estimated equation, estimated parameters, their standard errors, t values, p values and codes indicating their statistical significance, and some other statistics like standard error of the residuals or R^2 value of the equation.

10.4.2 User options of `systemfit`

The user can modify the default estimation method by providing additional optional arguments, e.g. to specify instrumental variables or to impose parameter restrictions. All optional arguments are described in the following:

10.4.2.1 Equation labels

The optional argument `eqnlabels` allows the user to label the equations. It has to be a vector of strings. If this argument is not provided, the labels are taken from the names of the equations in argument `eqns`. And if the equations have no names, they are numbered consecutively. Hence, the following command has the same effect as the command above.

```
> fitsur <- systemfit("SUR", list(eqDemand, eqSupply), eqnlabels = c("demand",
+     "supply"))
```

10.4.2.2 Instrumental variables

The instruments for a 2SLS, W2SLS or 3SLS estimation can be specified by the argument `inst`. If the same instruments should be used for all equations, `inst` must be a one-sided formula. If different instruments should be used for the equations, `inst` must be a list that contains a one-sided formula for each equation. The first example uses the same instruments for all equations, and the second uses different instruments:

```
> fit3sls <- systemfit("3SLS", list(demand = eqDemand, supply = eqSupply),
+     inst = ~income + farmPrice + trend)
> fit3sls2 <- systemfit("3SLS", list(demand = eqDemand, supply = eqSupply),
+     inst = list(~farmPrice + trend, ~income + farmPrice + trend))
```

10.4.2.3 Data

Having all data in the global environment or attached to the search path is often inconvenient. Therefore, a data frame `data` can be provided that contains the variables of the model. In the following example we do not need to attach the data frame `Kmenta` before calling systemfit:

```
> fitsur <- systemfit("SUR", list(eqDemand, eqSupply), data = Kmenta)
```

10.4.2.4 Parameter restrictions

As outlined in section 10.2.2, parameter restrictions can be imposed in two ways. The first way is to use the transformation matrix T that can be specified by the argument `TX`. If we want to impose the restriction, say $\beta_2 = -\beta_6$, we can do this as follows

```
> tx <- matrix(0, nrow = 7, ncol = 6)
> tx[1, 1] <- 1
> tx[2, 2] <- 1
> tx[3, 3] <- 1
> tx[4, 4] <- 1
> tx[5, 5] <- 1
> tx[6, 2] <- -1
> tx[7, 6] <- 1
> fitsurTx <- systemfit("SUR", list(eqDemand, eqSupply), TX = tx)
```

The first line creates a 7×6 matrix of zeros, where 7 is the number of unrestricted coefficients and 6 is the number of restricted coefficients. The following seven lines specify this matrix in a way that the unrestricted coefficients (β) are assigned to the restricted coefficients (β^*) with $\beta_1 = \beta_1^*$, $\beta_2 = \beta_2^*$, $\beta_3 = \beta_3^*$, $\beta_4 = \beta_4^*$, $\beta_5 = \beta_5^*$, $\beta_6 = -\beta_2^*$, and $\beta_7 = \beta_6^*$. Finally the model is estimated with restriction $\beta_2^* = \beta_2 = -\beta_6$ imposed.

The second way to impose parameter restrictions is to use the matrix R and the vector q (see section 10.2.2). Matrix R can be specified with the argument `R.restr` and vector q with argument `q.restr`. We convert the restriction specified above to $\beta_2 + \beta_6 = 0$ and impose it in the second way:

```
> Rmat <- matrix(0, nrow = 1, ncol = 7)
> Rmat[1, 2] <- 1
> Rmat[1, 6] <- 1
> qvec <- c(0)
> fitsurRmat <- systemfit("SUR", list(eqDemand, eqSupply), R.restr = Rmat,
+     q.restr = qvec)
```

The first line creates a 1×7 matrix of zeros, where 1 is the number of restrictions and 7 is the number of unrestricted coefficients. The following two lines specify this matrix in a way that the multiplication with the parameter vector results in $\beta_2 + \beta_6$. The fourth line creates a vector with a single element that contains the left hand side of the restriction, i.e. zero. Finally the model is estimated with restriction $\beta_2 + \beta_6 = 0$ imposed.

10.4.2.5 Iteration control

The estimation methods WLS, SUR, W2SLS and 3SLS need a covariance matrix of the residuals that can be calculated from a first-step OLS or 2SLS estimation (see section 10.2.3). If the argument `maxiter` is set to a number larger than one, this procedure is iterated and at each iteration the covariance matrix is calculated from the previous step estimation. This iteration is repeated until the maximum number of iterations is reached or the parameter estimates have converged. The maximum number of iterations is specified by the argument `maxiter`. Its default value is one, which means no iteration. The convergence criterion is

$$\sqrt{\frac{\sum_i (b_{i,g} - b_{i,g-1})^2}{\sum_i b_{i,g-1}^2}} < \texttt{tol} \tag{10.70}$$

where $b_{i,g}$ is the ith coefficient of the gth iteration. The default value of `tol` is 10^{-5}.

10.4.2.6 Residual covariance matrix

It was explained in section 10.2.3 that several different formulas have been proposed to calculate the residual covariance matrix. The user can specify, which formula **systemfit** should use. Possible values of the argument `rcovformula` are presented in table 10.1. By default, **systemfit** uses equation (10.57).

Table 10.1: Possible values of argument `rcovformula`

argument `rcovformula`	equation
0	10.56
1 or 'geomean'	10.57
2 or 'Theil'	10.59
3 or 'max'	10.58

Furthermore, the user can specify whether the means should be subtracted from the residuals before (10.56), (10.57), (10.58), or (10.59) are applied to calculate the residual covariance matrix (see section 10.2.3). The corresponding argument is called `centerResiduals`. It must be either "`TRUE`" (subtract the means) or "`FALSE`" (take the unmodified residuals). The default value of `centerResiduals` is "`FALSE`".

10.4.2.7 3SLS formula

As discussed in sections 10.2.1 and 10.2.2, there exist several different formulas to perform a 3SLS estimation. The user can specify the applied formula by the argument `formula3sls`. Possible values are presented in table 10.2. The default value is 'GLS'.

Table 10.2: Possible values of argument `formula3sls`

argument `formula3sls`	equation (unrestricted)	equation (restricted)
'GLS'	10.20	10.47
'IV'	10.23	10.49
'GMM'	10.25	10.51
'Schmidt'	10.27	10.53
'EViews'	10.29	10.55

10.4.2.8 Degrees of freedom for t tests

There exist two different approaches to determine the degrees of freedom to perform t tests on the estimated parameters (section 10.2.4). This can be specified with argument `probdfsys`. If it is `TRUE`, the degrees of freedom of the whole system are taken. In contrast, if `probdfsys` is `FALSE`, the degrees of freedom of the single equation are taken. By default, `probdfsys` is `TRUE`, if any restrictions are specified using either the argument `R.restr` or the argument `TX`, and it is `FALSE` otherwise.

10.4.2.9 Sigma squared

In case of OLS or 2SLS estimations, argument `single.eq.sigma` can be used to specify, whether different σ^2s for each single equation or the same σ^2 for all equations should be used. If argument `single.eq.sigma` is `TRUE`, equations (10.9) and (10.17) are applied. In contrast, if argument `single.eq.sigma` is `FALSE`, equations (10.8) and (10.16) are applied. By default, `single.eq.sigma` is `FALSE`, if any restrictions are specified using either the argument `R.restr` or the argument `TX`, and it is `TRUE` otherwise.

10.4.2.10 System options

Finally, two options regarding some internal calculations are available. First, argument `solvetol` specifies the tolerance level for detecting linear dependencies when inverting a matrix or calculating a determinant (using functions `solve` and `det`). The default value depends on the used computer system and is equal to the default tolerance level of `solve` and `det`. Second, argument `saveMemory` can be used in case of large data sets to accelerate the estimation by omitting some calculation that are not crucial for the basic estimation. Currently, only the calculation of McElroy's R^2 is omitted. The default value of argument `saveMemory` is `TRUE`, if the argument `data` is specified and the number of observations times the number of equations is larger than 1000, and it is `FALSE` otherwise.

10.4.3 The wrapper function `systemfitClassic`

The wrapper function `systemfitClassic` can be applied to (classical) panel-like data in long format[6] if the regressors are the same for all equations. This function is called by

```
> systemfitClassic(method, formula, eqnVar, timeVar, data)
```

The mandatory arguments are `method`, `formula`, `eqnVar`, and `timeVar`. Argument `method` is the same as in `systemfit` (see section 10.4.1). The second argument `formula` is a standard formula in R that will be applied to all equations. Argument `eqnVar` specifies the variable name indicating the equation to which the observation belongs, and argument `timeVar` specifies the variable name indicating the time. Finally, `data` is a `data.frame` that contains all required data.

We demonstrate the usage of `systemfitClassic` using an example taken from Theil (1971, pp. 295, 300) that is based on Grunfeld (1958). We want to estimate a model for gross investment of 2 US firms in the years 1935–1954:

$$\texttt{invest}_{it} = \beta_1 + \beta_2 * \texttt{value}_{it} + \beta_3 * \texttt{capital}_{it} \tag{10.71}$$

where `invest` is the gross investment of firm i in year t, `value` is the market value of the firm at the end of the previous year, and `capital` is the capital stock of the firm at the end of the previous year.

This model can be estimated by

[6] Panel data can be either in "long format" (different individuals are arranged below each other) or in "wide format" (different individuals are arranged next to each other).

```
> data("GrunfeldTheil")
> theilSur <- systemfitClassic("SUR", invest ~ value + capital,
+     "firm", "year", data = GrunfeldTheil)
```

The first line loads example data that are included with the package. And then, a seemingly unrelated regression is performed, where the variable "firm" indicates the firm and the variable "year" indicates the time.

The function `systemfitClassic` has also an optional argument `pooled` that is a logical variable indicating whether the coefficients should be restricted to be equal in all equations. By default, this argument is set to "`FALSE`". In addition all optional arguments of `systemfit` (see section 10.4.2) except for `eqnLabels` and `TX` can be used with `systemfitClassic`, too.

10.4.4 Testing linear restrictions

As described in section 10.2.6, linear restrictions can be tested by an F test, Wald test or LR test. The corresponding functions in package **systemfit** are `ftest.systemfit`, `waldtest.systemfit`, and `lrtest.systemfit`, respectively.

We will now test the restriction $\beta_2 = -\beta_6$ that was specified by the matrix `Rmat` and the vector `qvec` in the example above (p. 265).

```
> ftest.systemfit(fitsur, Rmat, qvec)

F-test for linear parameter restrictions in equation systems
F-statistic: 0.9322
degrees of freedom of the numerator: 1
degrees of freedom of the denominator: 33
p-value: 0.3413

> waldtest.systemfit(fitsur, Rmat, qvec)

Wald-test for linear parameter restrictions in equation systems
Wald-statistic: 0.6092
degrees of freedom: 1
p-value: 0.4351

> lrtest.systemfit(fitsurRmat, fitsur)

Likelihood-Ratio-test for parameter restrictions in equation systems
LR-statistic: 1.004
```

```
degrees of freedom: 1
p-value: 0.3163
```

The linear restrictions are tested by an F test first, then by a Wald test, and finally by an LR test. The functions `ftest.systemfit` and `waldtest.systemfit` have three arguments. The first argument must be an unrestricted regression returned by `systemfit`. The second and third argument are the restriction matrix R and the vector q as described in section 10.2.2. In contrast, the function `lrtest.systemfit` needs two arguments. The first argument must be a restricted and the second an unrestricted regression returned by `systemfit`.

All tests print a short explanation first. Then the empirical test statistic and the degree(s) of freedom are reported. Finally the p value is printed. While there is some variation of the p values across the three different tests, all tests suggest the same decision: The null hypothesis $\beta_2 = -\beta_6$ cannot be rejected at any reasonable level of significance.

10.4.5 Hausman test

A Hausman test, which is described in section 10.2.7, can be carried out with following commands:

```
> fit2sls <- systemfit("2SLS", list(demand = eqDemand, supply = eqSupply),
+     inst = ~income + farmPrice + trend, data = Kmenta)
> fit3sls <- systemfit("3SLS", list(demand = eqDemand, supply = eqSupply),
+     inst = ~income + farmPrice + trend, data = Kmenta)
> hausman.systemfit(fit2sls, fit3sls)

        Hausman specification test for consistency of the 3SLS estimation

data:  Kmenta
Hausman = 2.5357, df = 7, p-value = 0.9244
```

First of all, the model is estimated by 2SLS and then by 3SLS. Finally, in the last line the test is carried out by the command `hausman.systemfit`. This function requires two arguments: the result of a 2SLS estimation and the result of a 3SLS estimation. The Hausman test statistic is 2.536, which has a χ^2 distribution with 7 degrees of freedom under the null hypothesis. The corresponding p value is 0.924. This shows that the null hypothesis is not rejected at any reasonable level of significance. Hence, we can assume that the 3SLS estimator is consistent.

10.5 Testing reliability

In this section we test the reliability of the results from `systemfit` and `systemfitClassic`.

10.5.1 Kmenta (1986): Example on p. 685 (food market)

First, we reproduce an example taken from Kmenta (1986, p. 685). The data are available from Table 13-1 (p. 687), and the results are presented in Table 13-2 (p. 712) of this book.

Before starting the estimation, we load the data and specify the two formulas to estimate as well as the instrumental variables. Then the equation system is estimated by OLS, 2SLS, 3SLS, and iterated 3SLS. After each estimation the estimated coefficients are reported.

```
> data("Kmenta")
> eqDemand <- consump ~ price + income
> eqSupply <- consump ~ price + farmPrice + trend
> inst <- ~income + farmPrice + trend
> system <- list(demand = eqDemand, supply = eqSupply)
```

OLS estimation:

```
> fitOls <- systemfit("OLS", system, data = Kmenta)
> round(coef(summary(fitOls)), digits = 4)
```

```
                  Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)   99.8954     7.5194 13.2851   0.0000
eq 1 price         -0.3163     0.0907 -3.4882   0.0028
eq 1 income         0.3346     0.0454  7.3673   0.0000
eq 2 (Intercept)   58.2754    11.4629  5.0838   0.0001
eq 2 price          0.1604     0.0949  1.6901   0.1104
eq 2 farmPrice      0.2481     0.0462  5.3723   0.0001
eq 2 trend          0.2483     0.0975  2.5462   0.0216
```

2SLS estimation:

```
> fit2sls <- systemfit("2SLS", system, inst = inst, data = Kmenta)
> round(coef(summary(fit2sls)), digits = 4)
```

```
                  Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)   94.6333     7.9208 11.9474   0.0000
```

```
eq 1 price        -0.2436     0.0965 -2.5243   0.0218
eq 1 income        0.3140     0.0469  6.6887   0.0000
eq 2 (Intercept)  49.5324    12.0105  4.1241   0.0008
eq 2 price         0.2401     0.0999  2.4023   0.0288
eq 2 farmPrice     0.2556     0.0473  5.4096   0.0001
eq 2 trend         0.2529     0.0997  2.5380   0.0219
```

3SLS estimation:

```
> fit3sls <- systemfit("3SLS", system, inst = inst, data = Kmenta)
> round(coef(summary(fit3sls)), digits = 4)
                 Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)  94.6333     7.9208 11.9474   0.0000
eq 1 price        -0.2436     0.0965 -2.5243   0.0218
eq 1 income        0.3140     0.0469  6.6887   0.0000
eq 2 (Intercept)  52.1972    11.8934  4.3888   0.0005
eq 2 price         0.2286     0.0997  2.2934   0.0357
eq 2 farmPrice     0.2282     0.0440  5.1861   0.0001
eq 2 trend         0.3611     0.0729  4.9546   0.0001
```

Iterated 3SLS estimation:

```
> fitI3sls <- systemfit("3SLS", system, inst = inst, data = Kmenta,
+     maxit = 250)
> round(coef(summary(fitI3sls)), digits = 4)
                 Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)  94.6333     7.9208 11.9474   0.0000
eq 1 price        -0.2436     0.0965 -2.5243   0.0218
eq 1 income        0.3140     0.0469  6.6887   0.0000
eq 2 (Intercept)  52.6618    12.8051  4.1126   0.0008
eq 2 price         0.2266     0.1075  2.1086   0.0511
eq 2 farmPrice     0.2234     0.0468  4.7756   0.0002
eq 2 trend         0.3800     0.0720  5.2771   0.0001
```

The results above show that `systemfit` returns exactly the same coefficients and standard errors as published in Kmenta (1986, p. 712) except for two minor exemptions. Two standard errors of the 2SLS estimation deviate by 0.0001. However, this difference is likely due to rounding errors in `systemfit` or Kmenta (1986) and is so small that it empirically does not matter.

10.5.2 Greene (2003): Example 15.1 (Klein's model I)

Second, we try to replicate Klein's Model I (Klein, 1950) that is described in Greene (2003, p. 381). The data are available from the online complements to Greene (2003), Table F15.1 (http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm), and the estimation results are presented in Table 15.3 (p. 412).

Initially, the data are loaded and three equations as well as the instrumental variables are specified. As in the example before, the equation system is estimated by OLS, 2SLS, 3SLS, and iterated 3SLS, and estimated coefficients of each estimation are reported.

```
> data("KleinI")
> eqConsump <- consump ~ corpProf + corpProfLag + wages
> eqInvest <- invest ~ corpProf + corpProfLag + capitalLag
> eqPrivWage <- privWage ~ gnp + gnpLag + trend
> inst <- ~govExp + taxes + govWage + trend + capitalLag + corpProfLag +
+     gnpLag
> system <- list(Consumption = eqConsump, Investment = eqInvest,
+     "Private Wages" = eqPrivWage)
```

OLS estimation:

```
> kleinOls <- systemfit("OLS", system, data = KleinI)
> round(coef(summary(kleinOls)), digits = 3)
```

```
                 Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)   16.237      1.303  12.464    0.000
eq 1 corpProf       0.193      0.091   2.115    0.049
eq 1 corpProfLag    0.090      0.091   0.992    0.335
eq 1 wages          0.796      0.040  19.933    0.000
eq 2 (Intercept)   10.126      5.466   1.853    0.081
eq 2 corpProf       0.480      0.097   4.939    0.000
eq 2 corpProfLag    0.333      0.101   3.302    0.004
eq 2 capitalLag    -0.112      0.027  -4.183    0.001
eq 3 (Intercept)    1.497      1.270   1.179    0.255
eq 3 gnp            0.439      0.032  13.561    0.000
eq 3 gnpLag         0.146      0.037   3.904    0.001
eq 3 trend          0.130      0.032   4.082    0.001
```

2SLS estimation:

```
> klein2sls <- systemfit("2SLS", system, inst = inst, data = KleinI,
+     rcovformula = 0)
> round(coef(summary(klein2sls)), digits = 3)

                 Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)   16.555      1.321  12.534    0.000
eq 1 corpProf       0.017      0.118   0.147    0.885
eq 1 corpProfLag    0.216      0.107   2.016    0.060
eq 1 wages          0.810      0.040  20.129    0.000
eq 2 (Intercept)   20.278      7.543   2.688    0.016
eq 2 corpProf       0.150      0.173   0.867    0.398
eq 2 corpProfLag    0.616      0.163   3.784    0.001
eq 2 capitalLag    -0.158      0.036  -4.368    0.000
eq 3 (Intercept)    1.500      1.148   1.307    0.209
eq 3 gnp            0.439      0.036  12.316    0.000
eq 3 gnpLag         0.147      0.039   3.777    0.002
eq 3 trend          0.130      0.029   4.475    0.000
```

3SLS estimation:

```
> klein3sls <- systemfit("3SLS", system, inst = inst, data = KleinI,
+     rcovformula = 0)
> round(coef(summary(klein3sls)), digits = 3)

                 Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)   16.441      1.305  12.603    0.000
eq 1 corpProf       0.125      0.108   1.155    0.264
eq 1 corpProfLag    0.163      0.100   1.624    0.123
eq 1 wages          0.790      0.038  20.826    0.000
eq 2 (Intercept)   28.178      6.794   4.148    0.001
eq 2 corpProf      -0.013      0.162  -0.081    0.937
eq 2 corpProfLag    0.756      0.153   4.942    0.000
eq 2 capitalLag    -0.195      0.033  -5.990    0.000
eq 3 (Intercept)    1.797      1.116   1.611    0.126
eq 3 gnp            0.400      0.032  12.589    0.000
eq 3 gnpLag         0.181      0.034   5.307    0.000
eq 3 trend          0.150      0.028   5.358    0.000
```

iterated 3SLS estimation:

```
> kleinI3sls <- systemfit("3SLS", system, inst = inst, data = KleinI,
+     rcovformula = 0, maxit = 500)
> round(coef(summary(kleinI3sls)), digits = 3)
```

```
                 Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)   16.559      1.224  13.524    0.000
eq 1 corpProf       0.165      0.096   1.710    0.105
eq 1 corpProfLag    0.177      0.090   1.960    0.067
eq 1 wages          0.766      0.035  22.031    0.000
eq 2 (Intercept)   42.896     10.594   4.049    0.001
eq 2 corpProf      -0.357      0.260  -1.370    0.188
eq 2 corpProfLag    1.011      0.249   4.065    0.001
eq 2 capitalLag    -0.260      0.051  -5.115    0.000
eq 3 (Intercept)    2.625      1.196   2.195    0.042
eq 3 gnp            0.375      0.031  12.050    0.000
eq 3 gnpLag         0.194      0.032   5.977    0.000
eq 3 trend          0.168      0.029   5.805    0.000
```

Again, the results show that `systemfit` returns the same results as published in Greene (2003).[7] Also in this case we have two minor deviations, where only the last digit is different.

10.5.3 Theil (1971): Example on p. 295 (General Electric and Westinghouse)

Third, we estimate an example taken from Theil (1971, p. 295) that is based on Grunfeld (1958). The data are available from Table 7.1 (p. 296), and the results are presented on pages 295 and 300 of this book.

After loading the data and specifying the formula, the model is estimated by OLS and SUR. The coefficients of each estimation are reported.

```
> data("GrunfeldTheil")
> formulaGrunfeld <- invest ~ value + capital
```

OLS estimation (page 295)

[7] There are two typos in Table 15.3 (p. 412). Please take a look at the errata (http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm).

```
> theilOls <- systemfitClassic("OLS", formulaGrunfeld, "firm",
+     "year", data = GrunfeldTheil)
> round(coef(summary(theilOls)), digits = 3)

                             Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)               -9.956     31.374  -0.317    0.755
eq 1 value.General.Electric     0.027      0.016   1.706    0.106
eq 1 capital.General.Electric   0.152      0.026   5.902    0.000
eq 2 (Intercept)               -0.509      8.015  -0.064    0.950
eq 2 value.Westinghouse         0.053      0.016   3.368    0.004
eq 2 capital.Westinghouse       0.092      0.056   1.647    0.118
```

SUR estimation (page 300)

```
> theilSur <- systemfitClassic("SUR", formulaGrunfeld, "firm",
+     "year", data = GrunfeldTheil, rcovformula = 0)
> round(coef(summary(theilSur)), digits = 3)

                             Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)              -27.719     27.033  -1.025    0.320
eq 1 value.General.Electric     0.038      0.013   2.883    0.010
eq 1 capital.General.Electric   0.139      0.023   6.036    0.000
eq 2 (Intercept)               -1.252      6.956  -0.180    0.859
eq 2 value.Westinghouse         0.058      0.013   4.297    0.000
eq 2 capital.Westinghouse       0.064      0.049   1.308    0.208
```

The function `systemfitClassic`, which is a wrapper function to `systemfit` returns exactly the same results as published in Theil (1971, pp. 295, 300).

10.5.4 Greene (2003): Example 14.1 (Grunfeld's investment data)

Finally, we reproduce Example 14.1 of Greene (2003, p. 340) that is also based on Grunfeld (1958). The data are available from the online complements to Greene (2003), Table F13.1 (http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm), and the estimation results are presented in Tables 14.1 and 14.2 (p. 351).

First, we load the data and specify the formula to estimate. Then, the system is estimated by OLS, pooled OLS, SUR, and pooled SUR. Immediately after each estimation, the estimated coefficients are reported. Furthermore, the σ^2 values of the OLS estimations, and the residual covariance matrix as well as the residual correlation matrix of the SUR estimations are printed.

```
> data("GrunfeldGreene")
> formulaGrunfeld <- invest ~ value + capital
```

OLS estimation (Table 14.2):

```
> greeneOls <- systemfitClassic("OLS", formulaGrunfeld, "firm",
+     "year", data = GrunfeldGreene)
> round(coef(summary(greeneOls)), digits = 4)
```

```
                             Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)            -149.7825   105.8421 -1.4151   0.1751
eq 1 value.General.Motors      0.1193     0.0258  4.6172   0.0002
eq 1 capital.General.Motors    0.3714     0.0371 10.0193   0.0000
eq 2 (Intercept)              -6.1900    13.5065 -0.4583   0.6525
eq 2 value.Chrysler            0.0779     0.0200  3.9026   0.0011
eq 2 capital.Chrysler          0.3157     0.0288 10.9574   0.0000
eq 3 (Intercept)              -9.9563    31.3742 -0.3173   0.7548
eq 3 value.General.Electric    0.0266     0.0156  1.7057   0.1063
eq 3 capital.General.Electric  0.1517     0.0257  5.9015   0.0000
eq 4 (Intercept)              -0.5094     8.0153 -0.0636   0.9501
eq 4 value.Westinghouse        0.0529     0.0157  3.3677   0.0037
eq 4 capital.Westinghouse      0.0924     0.0561  1.6472   0.1179
eq 5 (Intercept)             -30.3685   157.0477 -0.1934   0.8490
eq 5 value.US.Steel            0.1566     0.0789  1.9848   0.0635
eq 5 capital.US.Steel          0.4239     0.1552  2.7308   0.0142
```

```
> round(sapply(greeneOls$eq, function(x) {
+     return(x$ssr/20)
+ }), digits = 3)
```

```
[1] 7160.294  149.872  660.829   88.662 8896.416
```

pooled OLS (Table 14.2):

```
> greeneOlsPooled <- systemfitClassic("OLS", formulaGrunfeld, "firm",
+     "year", data = GrunfeldGreene, pooled = TRUE)
> round(coef(summary(greeneOlsPooled$eq[[1]])), digits = 4)
```

```
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)           -48.0297    21.4802 -2.2360   0.0276
value.General.Motors    0.1051     0.0114  9.2360   0.0000
capital.General.Motors  0.3054     0.0435  7.0186   0.0000
```

```
> sum(sapply(greeneOlsPooled$eq, function(x) {
+     return(x$ssr)
+ }))/100
```

```
[1] 15708.84
```

SUR estimation (Table 14.1):

```
> greeneSur <- systemfitClassic("SUR", formulaGrunfeld, "firm",
+     "year", data = GrunfeldGreene, rcovformula = 0)
> round(coef(summary(greeneSur)), digits = 4)
```

```
                              Estimate Std. Error t value Pr(>|t|)
eq 1 (Intercept)             -162.3641    89.4592 -1.8150   0.0872
eq 1 value.General.Motors       0.1205     0.0216  5.5709   0.0000
eq 1 capital.General.Motors     0.3827     0.0328 11.6805   0.0000
eq 2 (Intercept)                0.5043    11.5128  0.0438   0.9656
eq 2 value.Chrysler             0.0695     0.0169  4.1157   0.0007
eq 2 capital.Chrysler           0.3085     0.0259 11.9297   0.0000
eq 3 (Intercept)              -22.4389    25.5186 -0.8793   0.3915
eq 3 value.General.Electric     0.0373     0.0123  3.0409   0.0074
eq 3 capital.General.Electric   0.1308     0.0220  5.9313   0.0000
eq 4 (Intercept)                1.0889     6.2588  0.1740   0.8639
eq 4 value.Westinghouse         0.0570     0.0114  5.0174   0.0001
eq 4 capital.Westinghouse       0.0415     0.0412  1.0074   0.3279
eq 5 (Intercept)               85.4233   111.8774  0.7635   0.4556
eq 5 value.US.Steel             0.1015     0.0548  1.8523   0.0814
eq 5 capital.US.Steel           0.4000     0.1278  3.1300   0.0061
```

```
> round(greeneSur$rcov, digits = 3)
```

```
                 General Motors Chrysler General Electric Westinghouse
General Motors         7216.044 -313.704          605.336      129.887
```

```
Chrysler             -313.704  152.849         2.047      16.661
General Electric      605.336    2.047       700.456     200.316
Westinghouse          129.887   16.661       200.316      94.912
US Steel            -2686.517  455.089      1224.405     652.716
                  US Steel
General Motors   -2686.517
Chrysler           455.089
General Electric  1224.405
Westinghouse       652.716
US Steel          9188.151
```

```
> round(summary(greeneSur)$rcor, digits = 3)
```

```
                 General Motors Chrysler General Electric Westinghouse US Steel
General Motors            1.000   -0.299            0.269        0.157   -0.330
Chrysler                 -0.299    1.000            0.006        0.138    0.384
General Electric          0.269    0.006            1.000        0.777    0.483
Westinghouse              0.157    0.138            0.777        1.000    0.699
US Steel                 -0.330    0.384            0.483        0.699    1.000
```

pooled SUR estimation (Table 14.1):

```
> greeneSurPooled <- systemfitClassic("WSUR", formulaGrunfeld,
+     "firm", "year", data = GrunfeldGreene, pooled = TRUE, rcovformula = 0)
> round(coef(summary(greeneSurPooled$eq[[1]])), digits = 4)
```

```
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)           -28.2467     4.8882 -5.7785        0
value.General.Motors    0.0891     0.0051 17.5663        0
capital.General.Motors  0.3340     0.0167 19.9859        0
```

```
> round(greeneSurPooled$rcov, digits = 3)
```

```
                 General Motors  Chrysler General Electric Westinghouse
General Motors        10050.525    -4.805        -7160.667    -1400.747
Chrysler                 -4.805   305.610        -1966.648     -123.920
General Electric      -7160.667 -1966.648        34556.603     4274.000
Westinghouse          -1400.747  -123.920         4274.000      833.357
```

```
US Steel              4439.989  2158.595       -28722.006    -2893.733
                    US Steel
General Motors      4439.989
Chrysler            2158.595
General Electric -28722.006
Westinghouse       -2893.733
US Steel           34468.976
```

```
> round(cov(residuals(greeneSurPooled)), digits = 3)
```

```
                 General Motors Chrysler General Electric Westinghouse
General Motors         9396.058 -432.370        -1322.236     -865.791
Chrysler               -432.370  167.403          174.038       89.336
General Electric      -1322.236  174.038         3733.260     1302.231
Westinghouse           -865.791   89.336         1302.231      564.156
US Steel               -897.950  260.431         -972.145     -180.385
                  US Steel
General Motors    -897.950
Chrysler           260.431
General Electric  -972.145
Westinghouse      -180.385
US Steel         10052.045
```

```
> round(summary(greeneSurPooled)$rcor, digits = 3)
```

```
                 General Motors Chrysler General Electric Westinghouse US Steel
General Motors            1.000   -0.345           -0.223       -0.376   -0.092
Chrysler                 -0.345    1.000            0.220        0.291    0.201
General Electric         -0.223    0.220            1.000        0.897   -0.159
Westinghouse             -0.376    0.291            0.897        1.000   -0.076
US Steel                 -0.092    0.201           -0.159       -0.076    1.000
```

For this example, the function `systemfitClassic` returns nearly the same results as published in Greene (2003).[8] Two different residual covariance matrices of the pooled SUR estimation are presented. The first is calculated without centering the results (see section 10.2.3).

[8]There are several typos and errors in Table 14.1 (p. 412). Please take a look at the errata of this book (http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm).

It is equal to the one published in the book (Greene, 2003, p. 351). The second residual covariance matrix is calculated after centering the results. It is equal to the one published in the errata (http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm).

10.6 Summary and outlook

In this article, we have described some of the basic features of the **systemfit** package for estimation of linear systems of equations. Many details of the estimation can be controlled by the user. Furthermore, the package provides some statistical tests for parameter restrictions and consistency of 3SLS estimation. It has been tested on a variety of datasets and has produced satisfactory for a few years. While the **systemfit** package performs the basic fitting methods, more sophisticated tools exist. We hope to implement missing functionalities in the near future.

Unbalanced datasets

Currently, the **systemfit** package requires that all equations have the same number of observations. However, many data sets have unbalanced observations.[9] Simply dropping data points that do not contain observations for all equations may reduce the number of observations considerably, and thus, the information utilized in the estimation. Hence, it is our intention to include the capability for estimations with unbalanced data sets as described in Schmidt (1977) in future releases of **systemfit**.

Serial correlation and heteroscedasticity

For all of the methods developed in the package, the disturbances of the individual equations are assumed to be independent and identically distributed (iid). The package could be enhanced by the inclusion of methods to fit equations with serially correlated and heteroscedastic disturbances (Parks, 1967).

Estimation methods

In the future, we wish to include more sophisticated estimation methods such as limited information maximum likelihood (LIML), full information maximum likelihood (FIML), generalized methods of moments (GMM) and spatial econometric methods.

[9] For instance, forestry datasets typically contain many observations of inexpensive variables (stem diameter, tree count) and few expensive variables such as stem height or volume.

Non-linear estimation

Finally, the **systemfit** package provides a function to estimate systems of non-linear estimations. However, the function `nlsystemfit` is currently under development and the results are not yet always reliable due to convergence difficulties.

Bibliography

Bates D (2004). "Least Squares Calculations in R." *R News*, **4**(1), 17–20. URL http://CRAN.R-project.org/doc/Rnews/.

Greene WH (2003). *Econometric Analysis*. Prentice Hall, 5th edition.

Greene WH (2006). "Information about SUR estimation in LIMDEP." Personal email on 2006/02/16.

Grunfeld Y (1958). *The Determinants of Corporate Investment*. Ph.D. thesis, University of Chicago.

Hausman JA (1978). "Specification Test in Econometrics." *Econometrica*, **46**, 1251–1272.

Judge GG, Griffiths WE, Hill RC, Lütkepohl H, Lee TC (1985). *The Theory and Practice of Econometrics*. John Wiley and Sons, 2nd edition.

Klein L (1950). *Economic Fluctuations in the United States, 1921–1941*. John Wiley, New York.

Kmenta J (1986). *Elements of econometrics*. Macmillan, New York, 2 edition.

McElroy MB (1977). "Goodness of Fit for Seemingly Unrelated Regressions." *Journal of Econometrics*, **6**, 381–387.

Parks RW (1967). "Effcient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated." *Journal of the American Statistical Association*, **62**, 500–509.

R Development Core Team (2005). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Schmidt P (1977). "Estimation of Seemingly Unrelated Regressions with Unequal Numbers of Observations." *Journal of Econometrics*, **5**, 365–377.

Schmidt P (1990). "Three-Stage Least Squares with Different Instruments for Different Equations." *Journal of Econometrics*, **43**, 389–394.

Theil H (1971). *Principles of Econometrics.* Wiley, New York.

Zellner A (1962). "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias." *Journal of the American Statistical Association*, **57**, 348–368.

Zellner A, Huang DS (1962). "Further Properties of Efficient Estimators for Seemingly Unrelated Regression Equations." *International Economic Review*, **3**(3), 300–313.

Zellner A, Theil H (1962). "Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations." *Econometrica*, **30**(1), 54–78.

Kapitel 11

Schlussbetrachtung

In dieser Arbeit wurden diverse Untersuchungen vorgestellt, die alle direkt oder indirekt die Wirkungsanalyse von Agrarpolitiken zum Ziel haben. In den empirischen Anwendungen der Agrarpolitikanalyse wurden unterschiedliche ökonomische Theorien und verschiedene quantitative Methoden verwendet. Dabei hängen die Anforderungen an die verwendeten Theorien und Methoden von der Komplexität der zu untersuchenden Agrarpolitiken, der betroffenen Wirtschaftsobjekte sowie deren Umfeld ab. Im Folgenden möchte ich nun kritisch diskutieren, in wie weit die von mir verwendeten Theorien und Methoden angemessen sind, und dabei insbesondere auf mögliche Schwächen, Unzulänglichkeiten und Verbesserungsmöglichkeiten eingehen.

An Empirical Investigation of the Demand for Bananas in Germany; Die EU Bananenmarktordnung und die Nachfrage in Deutschland

In den ersten beiden Aufsätzen wurde die Bananennachfrage und die Auswirkungen der EU Bananenmarktordnung mit einer relativ restriktiven ökonomischen Theorie untersucht. Erstens wurde nur die Nachfrageseite des Bananenmarktes betrachtet. Zweitens wurden perfekte Märkte angenommen und schließlich wurde angenommen, dass die Bananenmarktpolitik die Haushalte ausschließlich über den Bananenpreis beeinflusst. Diese vereinfachenden Annahmen sind allerdings in diesem speziellen Fall angemessen, da sie weitgehend der Realität entsprechen.

Auch die empirischen Methoden und dabei insbesondere die Verwendung der allgemeinen dynamischen Form des „Almost Ideal Demand Systems“ (AIDS) entsprechen den Anforderungen der Problemstellung. Allerdings würde ich heute nicht mehr eine lineare Approximation des AIDS (LA-AIDS), sondern das vollständige nicht-lineare AIDS mit einem linearen iterativen Verfahren (Browning und Meghir, 1991; Michalek und Keyzer, 1992; Blundell und Robin, 1999) ökonometrisch schätzen, da ich im achten Beitrag dieser Arbeit zeigen konnte, dass dieses Verfahren wesentlich genauere Schätzergebnisse als die lineare

Approximation liefert. Hinzu kommt, dass bei der Verwendung dieses vollständigen Nachfragesystems auch die Wohlfahrtseffekte der Bananenmarktordnung als „äquivalente Variation“ oder „kompensierende Variation“ hätten angegeben werden können.

Methodisches Vorgehen zur Modellierung der MTR-Beschlüsse; Economic Impact of the Mid-Term Review

Auch im dritten und vierten Beitrag wurde bei der Analyse der Auswirkungen der Agrarreform eine relativ restriktive ökonomische Theorie verwendet. Zum einen haben wir ausschließlich die Produktionsseite betrachtet und unterstellen den Landwirten, dass sie ausschließlich das Ziel der Gewinnmaximierung verfolgen. Weiterhin wurde angenommen, dass für die betrachteten Vorleistungen und Produktionsgüter vollkommene Märkte vorliegen, jedoch kein Land- bzw. Pachtmarkt existiert. Es ist davon auszugehen, dass einige dieser Annahmen in der Realität nicht immer erfüllt sind. In diesen Fällen haben wir versucht, mögliche Prognosefehler durch ad-hoc-Verfahren zu verringern. Beispielsweise haben wir die Opportunitätskosten der Familien-Arbeitskräfte unterhalb des zu erwartenden außerlandwirtschaftlichen Lohnsatzes gesetzt, um nicht von reiner Gewinnmaximierung auszugehen, sondern die Präferenzen für eine landwirtschaftliche Tätigkeit zu berücksichtigen.

Im Gegensatz zu den vorherigen Beiträgen ist die Wirkungsweise der untersuchten Politikänderung wesentlich vielschichtiger, denn es werden simultan diverse komplexe Maßnahmen (z.B. Entkopplung, Cross-Compliance, Milchpreissenkung, Einstellung der Roggenintervention) durchgeführt.

Die Methodik dieser Untersuchungen — die Erstellung des Gruppenhof- und Strukturwandelmodells sowie die Verwendung des CAPRI-Modells — sind im wesentlichen geeignete Instrumente, um die Forschungsziele dieser Analyse zu erreichen. Wie neuere Entwicklungen jedoch zeigen, hätte die Biogasproduktion als weiterer Produktionszweig im Gruppenhofmodell berücksichtigt werden müssen. Seit der Novellierung des Erneuerbare-Energien-Gesetzes (EEG) im August 2004 ist die Anzahl der Biogasanlagen in Deutschland sprunghaft angestiegen (Fachverband Biogas e.V., 2006b,a). Dadurch hat die Nachfrage nach Ackerflächen zur Erzeugung von Substraten stark zugenommen. Es ist daher davon auszugehen, dass der Rückgang des Ackerbaus und die Extensivierung des Grünlandes geringer ausfallen werden als mit unserem Gruppenhofmodell vorausgesagt wurde.

Analyse von Transaktionskosten auf dem ländlichen Kreditmarkt in Polen

Im Gegensatz zu den Annahmen der neoklassischen Mikroökonomie, die von transaktionskostenfreien Märkten ausgeht, können in der Realität häufig Transaktionskosten und Marktunvollkommenheiten beobachtet werden. Eine Nicht-Beachtung der Unvollkommenheiten kann daher in vielen Fällen zu falschen Ergebnissen und somit auch zu falschen Politikempfehlungen führen. Daher wird im fünften Beitrag die ökonomische Theorie um Transaktionskosten erweitert. Es wird folglich nicht mehr von vollkommenen Märkten ausgegangen, sondern es werden Marktunvollkommenheiten explizit beachtet. Des Weiteren wird auch Risikoaversion auf Seiten der Landwirte berücksichtigt. Für die hier vorgenommene Untersuchung — Transaktionskosten auf dem ländlichen Kreditmarkt in Polen — sind die getroffenen Annahmen angemessen. Allerdings könnte die Risikoaversion, die als multiplikativer Aufschlag auf den Zinssatz angenommen wurde, sicherlich noch realitätsnäher berücksichtigt werden. Beispielsweise würde eine empirische Analyse der Risikoaversion der betrachteten Landwirte eine deutlich genauere Aufteilung der beobachteten Zinsdifferenzen auf die möglichen Ursachen Risikoaversion und Transaktionskosten ermöglichen.

Ein weiterer Punkt ist, dass die Konkavität der geschätzten Produktionsfunktion erzwungen wurde, um die ermittelten Grenzprodukte des Kapitals aus einer mikroökonomisch konsistenten Technologie ableiten zu können. Jedoch ist die Konkavität der Produktionsfunktion zwar (zusammen mit Monotonie) eine hinreichende, aber keine notwendige Bedingung für mikroökonomische Konsistenz, denn es reicht aus, dass die Produktionsfunktion quasikonkav ist. Daher haben wir in späteren Anwendungen der hier vorgestellten Untersuchungsmethode nur die Quasikonkavität der Produktionsfunktion erzwungen.

Modeling Farm Households' Price Responses; Taxation of the Farm Household and Imperfect Labor Markets

Im sechsten Beitrag wird nun der Komplexitätsgrad der ökonomischen Theorie weiter erhöht, indem Elemente aus den bisherigen Beiträgen zusammengeführt werden. Zum einen wird mit einem Unternehmens-Haushalts-Modell (UHM) sowohl die Produktions- als auch die Konsumseite betrachtet. Zum anderen werden wie im fünften Beitrag zusätzlich Transaktionskosten und Marktunvollkommenheiten berücksichtigt.

Im siebten Beitrag wird schließlich die ökonomische Theorie nochmals weiterentwickelt, um eine spezifische Analyse verschiedener Steuerpolitiken zu ermöglichen. Somit können mit diesem erweiterten Unternehmens-Haushalts-Modell die Auswirkungen von diversen Standard-Steuern sowie speziellen landwirtschaftlichen Steuern unter gleichzeitiger Berücksichtigung von Transaktionskosten sowie der Produktions- und Konsumseite untersucht werden.

Diese erweiterte ökonomische Theorie ist gerade für die Agrarpolitikanalyse in Transformations- und Entwicklungsländern sehr geeignet. In vielen dieser Länder gibt es viele kleine Familienbetriebe, in denen die Produktions- und Konsumentscheidungen beispielsweise durch

Subsistenzwirtschaft eng verzahnt sind. Auch sind in diesen Ländern die institutionellen Rahmenbedingungen häufig noch nicht ausreichend entwickelt, sodass relativ hohe Transaktionskosten vorliegen können. Insbesondere auf Arbeitsmärkten lassen sich in der Regel hohe Transaktionskosten und Heterogenitäten beobachten. Daher konzentriert sich die hier vorliegende Untersuchung auf die Auswirkungen von Unvollkommenheiten auf den Märkten für landwirtschaftliche und außerlandwirtschaftliche Arbeit.

Da aber auch auf anderen Märkten relevante Transaktionskosten existieren können, bietet sich für die Zukunft eine Erweiterung des Ansatzes an, bei dem Unvollkommenheiten auf mehreren Gütermärkten berücksichtigt werden. Einen weiteren interessanten Forschungsaspekt stellt eine genauere Analyse der Transaktionskosten und deren Determinanten dar.

Im Rahmen dieser Untersuchung wurden fortgeschrittene ökonometrische Methoden angewendet und teilweise auch selbst entwickelt, sodass das oben beschriebene theoretische Unternehmens-Haushalts-Modell konsistent ökonometrisch geschätzt werden konnte.

Bei den ökonometrischen Schätzungen der Arbeitsmarktfunktionen konnte für viele erklärende Variablen kein statistisch signifikanter Einfluss festgestellt werden (siehe Tabelle 7.4). Für die hier vorliegende Untersuchung bedeutet die mangelnde Signifikanz der geschätzten Parameter jedoch kein zu großes Problem, denn dieser Beitrag ist auf die Anpassungsreaktionen landwirtschaftlicher Familienbetriebe fokussiert. Gleichwohl stellen die Auswirkungen verschiedener Einflussfaktoren auf das Arbeitsangebot, die Arbeitsnachfrage sowie die Lohnhöhe auf ländlichen Arbeitsmärkten eine interessante Forschungsfrage dar, sodass sich für die Zukunft eine detailliertere Untersuchung auf diesem Gebiet anbietet.

In dieser Untersuchung wurde die von der mikroökonomischen Theorie geforderte Konvexität der Profitfunktionen durch eine zweistufige Minimum-Distance Schätzung sichergestellt. Da diese Schätzmethode keine direkte Berechnung der Varianz-Kovarianz-Matrix der geschätzten Koeffizienten erlaubt, wurde diese mit einem Bootstrap-Verfahren ermittelt. Allerdings hat Andrews (2000) gezeigt, dass bei Ungleichheits-Restriktionen (wie z.B. bei Konvexitätsrestriktionen) der Bootstrap-Mechanismus zu inkonsistenten Ergebnissen führen kann. Eine Lösung hierfür bietet die Bayesianische Ökonometrie mit einer „Markov Chain Monte Carlo“ (MCMC) Simulation. Diese Methode stellt auch bei beschränkten Schätzungen (z.B. Profitfunktionen unter Konvexitätsrestriktionen) die stochastische Verteilung der geschätzten Koeffizienten zur Verfügung. Daher stellt das MCMC-Verfahren eine relevante Verbesserungsmöglichkeit der hier verwendeten Schätzstrategie dar.

How to Estimate the "Almost Ideal Demand System"

Das „Almost Ideal Demand System" (AIDS) ist ein in sehr vielen empirischen Nachfrageanalysen verwendetes Nachfragesystem und wurde auch im ersten, zweiten, sechsten und siebten Beitrag dieser Arbeit eingesetzt. Da sehr viele verschiedene Ansätze zur ökonometrischen Schätzung des AIDS bzw. einer linearen Approximation des AIDS existieren, standen wir vor der Frage, welche Schätzmethode wir in unseren empirischen Anwendungen verwenden sollten. Mit der hier vorliegenden Analyse konnten wir diese Frage nun beantworten und haben den „Iterated Linear Least Squares Estimator" (ILLE) ausgewählt.

Diese iterative lineare Schätzmethode wurde in diesem Aufsatz Michalek und Keyzer (1992) zugeschrieben. Allerdings wurde sie nach meinem derzeitigen Wissen zuerst von Browning und Meghir (1991) veröffentlicht. Blundell und Robin (1999) haben später die Konvergenz dieser Methode bewiesen und ihr den Namen „Iterated Linear Least Squares Estimator" (ILLE) gegeben.

Die Ergebnisse der Monte-Carlo-Simulation, mit der verschiedene Verfahren zur ökonometrischen Schätzung des „Almost Ideal Demand Systems" (AIDS) verglichen wurden, sind bisher noch nicht vollständig ausgewertet worden. So werden z.B. in den Tabellen 9.1 und 9.2 nur die durchschnittlichen Ergebnisse über alle 48 verschiedenen Rahmenbedingungen angegeben. Eine systematische Analyse der unterschiedlichen Schätzverfahren unter den verschiedenen Rahmenbedingungen kann voraussichtlich aufzeigen, in welchen Situationen bestimmte Verfahren besonders gute oder besonders schlechte Regressionsergebnisse liefern. Diese detaillierten Ergebnisse können in der angewandten Nachfrageanalyse sehr hilfreich sein.

systemfit: Simultaneous Equation Systems in R

Für die empirischen Anwendungen in dieser Arbeit brauchten wir eine Software, mit der die ökonometrischen Schätzungen, die statistischen Verfahren und diverse weitere Berechnungen durchgeführt werden konnten. Um dabei nicht zwischen verschiedenen Softwarepaketen wechseln zu müssen, sollte dieses Softwarepaket möglichst sämtliche benötigten Methoden bereitstellen. Die Wahl fiel auf die Statistik-Umgebung „R", da diese als Open-Source-Software an eigene Bedürfnisse angepasst werden kann. Allerdings waren die Verfahren zur ökonometrischen Schätzung von simultanen Gleichungssystemen, die wir zur Schätzung von Nachfragesystemen und Profitfunktionen benötigten, in „R" noch nicht sehr weit entwickelt. Daher habe ich das Zusatzpaket „systemfit" so weiterentwickelt, dass es jetzt sämtliche von uns benötigten Schätzverfahren zur Verfügung stellt.

Im Vergleich zu anderer Software hat es mehrere Vorzüge, aber auch einige Nachteile. So können zur Zeit noch keine Gleichungssysteme geschätzt werden, bei denen die einzelnen

Gleichungen eine unterschiedliche Anzahl von Beobachtungen aufweisen. Weiterhin bietet „systemfit“ bisher noch nicht die ökonometrischen Schätzmethoden „limited information maximum likelihood“ (LIML), „full information maximum likelihood“ (FIML) und „generalized methods of moments“ (GMM) an. Schließlich befindet sich die Funktion „nlsystemfit“ zur ökonometrischen Schätzung nicht-linearer Gleichungssysteme im Moment noch in der Entwicklung, sodass deren Ergebnisse nicht immer verlässlich sind. Diese Schwächen sollen in zukünftigen Versionen von „systemfit“ behoben werden. Darüber hinaus ist geplant, die zweistufige Schätzung von zensierten Mehrgleichungsmodellen (Shonkwiler und Yen, 1999) in „systemfit“ zu implementieren.

Literaturverzeichnis

Andrews, D. W. K. (2000): Inconsistency of the Bootstrap when a Parameter is on the Boundary of the Parameter Space. In: Econometrica, 68:399–405.

Blundell, R. W. und J. M. Robin (1999): Estimation in Large and Disaggregated Demand Systems: An Estimator for Conditionally Linear Systems. In: Journal of Applied Econometrics, 14:209–232.

Browning, M. und C. Meghir (1991): The Effects of Male and Female Labor Supply on Commodity Demands. In: Econometrica, 59:925–951.

Fachverband Biogas e.V. (2006a): Biogas in Deutschland 2005. In: http://www.biogas.org/datenbank/file/notmember/presse/Pressegespr_Hintergrunddaten1.pdf.

Fachverband Biogas e.V. (2006b): Biogas weiter im Aufwind. In: http://www.biogas.org/datenbank/file/notmember/presse/PM_JahresTagung060125.pdf.

Michalek, J. und M. A. Keyzer (1992): Estimation of a Two-stage LES-AIDS Consumer Demand System for Eight EC Countries. In: European Review of Agricultural Economics, 19(2):137–163.

Shonkwiler, J. S. und S. T. Yen (1999): Two-Step Estimation of a Censored System of Equations. In: American Journal of Agricultural Economics, 81:972–982.

Kapitel 12

Summary

In many countries the state heavily intervenes in the agricultural sector. For instance, domestic prices of agricultural products are regulated by import quotas, export subsidies and intervention prices. Furthermore, coupled or decoupled payments as well as several other subsidies (e.g. for credits) are granted and special agricultural taxes are raised. The analysis of the impacts of these interventions is complex and requires a detailed study of the supply and demand in the agricultural sector. One main focus of the research in agricultural economics is the analysis of the behavior of consumers and agricultural producers. Based on these results the impacts of agricultural policies can be determined.

This dissertation consists of nine papers that all deal with the analysis of supply and demand in the agricultural sector. They cover microeconomic theory as well as econometric methods and programming models. Furthermore, they contain empirical applications of the theory and of the quantitative methods.

An Empirical Investigation of the Demand for Bananas in Germany

The first paper investigates the demand for bananas of German households. Monthly household survey data of three different household types are analyzed for the period 1986–1998. The econometric estimation is based on log-linear demand functions and on a general dynamic version of the "Almost Ideal Demand System" (Deaton and Muellbauer, 1980a,b; Anderson and Blundell, 1982). Demand for bananas is characterized by significant habit persistence. It is significantly responsive to own price as well as income changes. Furthermore, there is evidence that other categories of fruit are both gross and net substitutes for bananas.

The EU banana regime and the demand in Germany

The implementation of the EU banana regime in 1993 has increased the consumer price of bananas in Germany by approximately 38%. Its impacts on the demand for fresh fruit of

German households is analyzed by using the results of the demand analysis presented in the previous paper. The policy-induced price increase of bananas has reduced banana consumption of the examined household types by 12.5% to 17.6% and increased the consumption of substitutes, namely apples and pears and the category “other types of fresh fruit”. The households' expenditure on total fresh fruit increased by 4.7% to 6.1%.

Methodology to Model the Impact of the Mid-Term Review

This paper describes the methodology that was used to analyze the impact of the latest reform of the EU Common Agricultural Policy — the Mid-Term Review (MTR) — on the agricultural sector in Schleswig-Holstein, Mecklenburg-Vorpommern (two federal states in Germany), Germany and Europe as well as on the world wheat market. A detailed farm group model was developed to model agricultural production in Schleswig-Holstein and Mecklenburg-Vorpommern. This model provides very disaggregated results for different farm sizes, farm types and subregions of these federal states. Furthermore, the existing CAPRI model was used to model the impact on agricultural production and markets on the national, European, and global level.

Economic Impact of the Mid-Term Review

This study analyzes the impact of the latest reform of the EU Common Agricultural Policy — the Mid-Term Review (MTR) — on the agricultural sector in Schleswig-Holstein. First, a very detailed farm group linear programming model is built to quantify the effects on agricultural production and farm incomes. The production adjustment to the MTR and its impact on farm profit vary significantly between individual farms. These results depend mainly on the farm type and the resource endowments of the farms. Second, the impact on structural change is examined with a farm survival model. Although the MTR clearly reduces the incomes of several farm types, it accelerates the structural change only gradually.

An Investigation of Transaction Costs on Rural Credit Markets in Poland

The functioning of the rural credit market is an essential prerequisite for the competitiveness of agricultural production. There exists a controversial discussion whether the low investment rate that can be observed in many transition countries is a consequence of limited access to rural credit markets. Therefore, we develop a methodology to quantify the impact of transaction costs on rural credit markets. We derive the return on agricultural investment from an econometrically estimated production function. After correcting for risk aversion, the gap between these internal returns and observed interest rates is used as a proxy for transaction costs. An empirical application of this method shows that the low investment rates of Polish farms can be attributed to high transaction costs on the rural credit market in Poland.

Modeling Farm Households' Price Responses

We develop a farm household model to analyze price responses of farm households. This model incorporates various types of transaction costs as well as labor heterogeneity. Non-proportional variable transaction costs or labor heterogeneity imply that production and consumption decisions become non-separable, even when the household buys or sells labor. An empirical model is estimated using data from Mid-West Poland. The results show that non-proportional variable transaction costs and labor heterogeneity significantly influence household behavior. Not all price elasticities, however, change significantly if these are neglected.

Taxation of the Farm Household and Imperfect Labor Markets

Recent analyses of agricultural tax policies show how standard presumptions in public finance literature change when market imperfections are taken into account (Hoff et al., 1993). This paper provides a theoretical analysis as well as an econometric estimation of an interdependent farm household model (FHM) approach considering imperfect labor markets and various tax policy instruments. Comparative static analysis supports the results of recent studies showing that neither standard nor land taxes are theoretically superior when compared to agricultural tax instruments. However, empirical estimation with individual household data from Poland partly confirms standard presumptions, i.e. even when markets are imperfect, income and consumption taxes imply negligible production adjustments.

How to Estimate the "Almost Ideal Demand System"

The "Almost Ideal Demand System" (AIDS) (Deaton and Muellbauer, 1980a) is widely used in applied demand analysis. To avoid complicated non-linear estimation of the original AIDS, most empirical studies use a linear approximation of AIDS (LA-AIDS). However, this leads to several econometric and theoretical problems. Several scholars have suggested approaches to circumvent these problems, but none of these approaches solves all of them. In this framework the paper presents the results of a Monte Carlo analysis of various AIDS and LA-AIDS approaches. The main result is that the non-linear AIDS can be very accurately estimated by an iterative linear estimation procedure (Browning and Meghir, 1991; Michalek and Keyzer, 1992; Blundell and Robin, 1999).

systemfit: Simultaneous Equation Systems in R

This paper presents the software package "systemfit". It provides the capability to estimate systems of linear equations within the "R" (R Development Core Team, 2005) programming environment. Many details of the estimation can be controlled by the user. Furthermore, the package provides some statistical tests for parameter restrictions and consistency of 3SLS estimation. It has been tested on a variety of datasets and has produced reliable results in the last years.

Bibliography

Anderson, G. J. and R. W. Blundell (1982): Estimation and Hypothesis Testing in Dynamic Singular Equation Systems. In: Econometrica, 50(6):1559–1571.

Blundell, R. W. and J. M. Robin (1999): Estimation in Large and Disaggregated Demand Systems: An Estimator for Conditionally Linear Systems. In: Journal of Applied Econometrics, 14:209–232.

Browning, M. and C. Meghir (1991): The Effects of Male and Female Labor Supply on Commodity Demands. In: Econometrica, 59:925–951.

Deaton, A. and J. Muellbauer (1980a): An Almost Ideal Demand System. In: The American Economic Review, 70:312–326.

Deaton, A. and J. Muellbauer (1980b): Economics and Consumer Behaviour. Cambridge University Press, Cambridge.

Hoff, K., A. Braverman and J. E. Stiglitz (Eds.) (1993): The Economics of Rural Organization. Oxford University Press, New York.

Michalek, J. and M. A. Keyzer (1992): Estimation of a Two-stage LES-AIDS Consumer Demand System for Eight EC Countries. In: European Review of Agricultural Economics, 19(2):137–163.

R Development Core Team (2005): R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. In: http://www.R-project.org. ISBN 3-900051-07-0.

Lebenslauf

Arne Henningsen, geboren am 30. Januar 1974 in Flensburg

Schule

1980 – 1984	Grundschule Steinbergkirche
1984 – 1993	Klaus-Harms-Schule Kappeln, Gymnasium

Zivildienst

1993 – 1994	Klappschau, gemeinnützige Gesellschaft für Seelenpflege mbH, Schleswig

Berufsausbildung zum Landwirt

1994 – 1995	Landwirtschaftlicher Ausbildungsbetrieb J. Böckmann, Kolauerhof
1995 – 1996	Versuchsgut Hohenschulen, Achterwehr

Universität

1996 – 2001	Diplom-Agraringenieur, Fachrichtung Wirtschafts- und Sozialwissenschaften des Landbaus, Christian-Albrechts-Universität zu Kiel
1999 – 2000	European Master in Agricultural Economics and Agribusiness, Wageningen Agricultural University, Niederlande
2001 – heute	Wissenschaftlicher Mitarbeiter am Institut für Agrarökonomie, Christian-Albrechts-Universität zu Kiel