

ANSÄTZE ZUR (AUF-)LÖSUNG EINES ALTEN METHODENSTREITS: ÖKONOMETRISCHE SPEZIFIKATION VON PROGRAMMIERUNGSMODELLEN ZUR AGRARANGEBOTSANALYSE

Referat

von

Thomas Heckeley und Hendrik Wolff*

*Institut für Agrarpolitik, Marktforschung und Wirtschaftssoziologie der Universität Bonn
Nußallee 21, 53115 Bonn; e-mail: heckeley@agp.uni-bonn.de, wolff@agp.uni-bonn.de

41. Jahrestagung der
Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V.
vom 8. Bis 10. Oktober 2001 in Braunschweig

ANSÄTZE ZUR (AUF-)LÖSUNG EINES ALTEN METHODENSTREITS: ÖKONOMETRISCHE SPEZIFIKATION VON PROGRAMMIERUNGSMODELLEN ZUR AGRARANGEBOTSANALYSE

von

Thomas Heckelei und Hendrik Wolff¹

1 Einleitung

Die quantitative Modellierung des Angebotsverhaltens von Landwirten wird schon seit Jahrzehnten durch das Nebeneinander von zwei unterschiedlichen methodischen Vorgehensweisen gekennzeichnet: zum Einen die ökonomische Spezifikation von Verhaltensfunktionen auf Basis beobachteter Daten, zum Anderen die Verwendung von Programmierungsmodellen mit expliziter Optimierung einer ökonomischen Zielfunktion unter Beachtung technologischer Nebenbedingungen. Diese ‚Koexistenz‘ ergab sich aus den spezifischen Vor- und Nachteilen der jeweiligen Methode für bestimmte Analysezwecke, an der sich auch durch die vielfältigen Entwicklungen auf beiden Seiten bisher nichts grundsätzlich geändert hat: Ökonometrische Modelle erlauben eine empirische Überprüfung ökonomischer Hypothesen auf Basis wohldefinierter statistischer Verfahren. Der Vorteil dieser eingebauten empirischen Modellvalidierung ist aber häufig mit dem Zwang zur erheblichen Aggregation und Vereinfachung der Modellstruktur verbunden. Programmierungsmodelle hingegen erlauben eine differenzierte Abbildung der landwirtschaftlichen Technologie in Form von Nebenbedingungen. Die allermeisten angebotsrelevanten agrarpolitischen Instrumente sowie Ressourcenbeschränkungen können in der Simulation direkt abgebildet werden und die häufig übliche aktivitätsanalytische Differenzierung der Produktionstechnologie erlaubt einen direkten Bezug zu betriebswirtschaftlichen Konzepten. Der Nachteil der Programmierungsmodelle liegt vor allem in der schwierigen Einbeziehung beobachteten Verhaltens bei der Modellspezifikation, d.h. die Nutzung von Beobachtungen zu modellendogenen Variablen.

Die Positiv Mathematische Programmierung (PMP) nach HOWITT (1995a) versprach Abhilfe in dieser Hinsicht. Sie erlaubt die exakte Kalibrierung eines Programmierungsmodells auf die beobachteten Aktivitätsumfang eines Basisjahres durch Einführung zusätzlicher nicht-linearer Terme in der Zielfunktion. Diese Eigenschaft im Zusammenhang mit einer weniger sprunghaften Angebotsreaktion im Vergleich zu LP Spezifikationen ließen PMP-Ansätze zur vorherrschenden Praxis bei der Spezifikation aggregierter Programmierungsmodelle werden (siehe z.B. Literaturliste in HOWITT 1995a und CYPRIIS 2000). PARIS und HOWITT (1998) schufen die Grundlage für die Einbeziehung mehrerer Beobachtungen durch Einführung eines ökonomischen Kriteriums, was dann schließlich auch zur ersten „Schätzung von Programmierungsmodellen“ mit Querschnittsdaten führte (HECKELEI & BRITZ 2001).

Dieser Beitrag beschäftigt sich mit den Möglichkeiten der Schätzung von Programmierungsmodellen. Zunächst soll aufgezeigt werden, warum PMP keine sinnvolle Grundlage für die konsistente Schätzung der Parameter bietet. Statt dessen wird eine allgemein anwendbare Alternative vorgestellt. In den daran anschließenden Abschnitten wird die grundsätzliche Methode exemplarisch an drei verschiedenen Gewinnmaximierungsmodellen illustriert, die den Bogen von einem typischen PMP-modell über CES-Produktionsfunktionen in Verbindung mit mehreren fixen Faktoren zu neueren dualitätstheoretischen Angebotsmodellen mit expliziter Landallokation schlagen. Die Darstellungen zeigen auch potentielle Vorteile der Methode gegenüber klassischer Schätzung von Verhaltensfunktionen und streifen ebenfalls die Möglichkeiten der Einbeziehung von a-priori Informationen im Rahmen der „Generalised Maximum Entropy“ (GME) Methode. Der letzte Abschnitt zieht Schlußfolgerungen und weist auf weiteren Forschungsbedarf hin.

¹ Thomas Heckelei, Hendrik Wolff, Institut für Agrarpolitik, Marktforschung und Wirtschaftssoziologie der Universität Bonn, Nußallee 21, 53332 Bonn; email: heckelei@agp.uni-bonn.de; wolff@agp.uni-bonn.de

2 PMP - Kritik und eine Allgemeine Alternative

Die grundsätzliche Idee von PMP besteht in der Nutzung der Dualwerte von Kalibrierungsbeschränkungen, die das Optimierungsproblem auf beobachtete Aktivitätsumfänge zwingt (Schritt 1). Diese Dualwerte werden zur Spezifikation nicht-linearer Terme in der Zielfunktion eingesetzt, die eine exakte Reproduktion der beobachteten Werte ohne Kalibrierungsbeschränkungen erlauben (Schritt 2). Ausgehend von einem typischen LP zur landwirtschaftlichen Angebotsanalyse kann Schritt 1 dargestellt werden als

$$(1) \quad \begin{aligned} \max_{\mathbf{l}} Z &= \mathbf{p}'\mathbf{l} - \mathbf{c}'\mathbf{l} \quad \text{so dass} \\ \mathbf{A}\mathbf{l} &\leq \mathbf{b} \quad [\boldsymbol{\lambda}], \quad \mathbf{l} \leq (\mathbf{l}^0 + \boldsymbol{\varepsilon}) \quad [\boldsymbol{\rho}], \quad \mathbf{l} \geq \mathbf{0} \end{aligned}$$

wobei Z der Zielfunktionswert, \mathbf{p} der $(N \times 1)$ Produktpreisvektor, \mathbf{l} der $(N \times 1)$ Vektor von nicht-negativen Aktivitätsumfängen, \mathbf{c} der $(N \times 1)$ Vektor variabler Kosten pro Aktivitätseinheit, \mathbf{A} die $(M \times N)$ Koeffizientenmatrix, \mathbf{b} der $(M \times 1)$ Vektor der Ressourcenbeschränkungen, $\boldsymbol{\lambda}$ der $(M \times 1)$ Schattenpreisvektor der Ressourcen, \mathbf{l}^0 der $(N \times 1)$ Vektor beobachteter Aktivitätsumfänge, wobei der Index „o“ für die beobachteten Variablen steht, $\boldsymbol{\varepsilon}$ ein $(N \times 1)$ Vektor kleiner Zahlen und $\boldsymbol{\rho}$ der $(N \times 1)$ Dualvariablenvektor der Kalibrierungsbeschränkungen. Im zweiten Schritt von PMP werden die Dualwerte $\boldsymbol{\rho}$ zur Spezifikation einer nichtlinearen variablen Kostenfunktion $K^V(\mathbf{l}^0)$ herangezogen, so dass die „variablen“ Grenzkosten $\mathbf{GK}^V(\mathbf{l}^0)$ der Aktivitäten gleich der Summe aus den bekannten Kosten \mathbf{c} und den „nicht spezifizierten Grenzkosten“ $\boldsymbol{\rho}$ sind. Im Falle der häufig verwendeten quadratischen Funktionsform ergibt sich demnach die folgende Kalibrierungsbedingung:

$$(2) \quad \mathbf{GK}^V = \frac{\partial K^V(\mathbf{l}^0)}{\partial \mathbf{l}} = \mathbf{d} + \mathbf{Q}\mathbf{l}^0 = \mathbf{c} + \boldsymbol{\rho},$$

wobei \mathbf{d} einen $(N \times 1)$ Vektor und \mathbf{Q} eine $(N \times N)$ symmetrische, positiv definite Matrix von Parametern darstellen, die jeweils mit dem linearen und quadratischen Termen der variablen Kostenfunktion korrespondieren. Die Kalibrierungsbedingung bezieht nicht die Opportunitätskosten der Nutzung fixer Ressourcen mit ein, da diese im resultierenden Modell

$$(3) \quad \begin{aligned} \max_{\mathbf{l}} Z &= \mathbf{p}'\mathbf{l} - \mathbf{d}'\mathbf{l} - 0,5\mathbf{l}'\mathbf{Q}\mathbf{l} \quad \text{so dass} \\ \mathbf{A}\mathbf{l} &\leq \mathbf{b} \quad [\boldsymbol{\lambda}], \quad \mathbf{l} \geq \mathbf{0} \end{aligned}$$

weiterhin durch die Ressourcenbeschränkungen eingeführt werden. Um das unterbestimmte Gleichungssystem (2) mit $N+N(N+1)/2$ Parametern und N Gleichungen zu lösen, wurden in der Literatur eine Vielfalt von Ansätzen von einfachen Rechenregeln mit a-priori Festlegung von Parametern (HOWITT, 1995a) über die Einbeziehung von Angebotselastizitäten (HELMING et al. 2001) bis hin zum Maximum Entropy Kriterium vorgeschlagen (PARIS & HOWITT, 1998). Solange die Bedingung (2) eingehalten wird, ist eine Kalibrierung des resultierenden Optimierungsmodells garantiert, jedoch die verschiedenen Möglichkeiten der Ausprägung von \mathbf{d} und \mathbf{Q} implizieren z.T. drastische Unterschiede im Hinblick auf das Simulationsverhalten (siehe CYPRI 2000 oder HECKELEI & BRITZ 2001). In diesem Beitrag soll aber nicht die Kalibrierung, sondern die *Schätzung* von Programmierungsmodellen im Vordergrund stehen. PARIS & HOWITT (1998 und 2001) weisen bereits auf die Möglichkeit hin, mehrere Beobachtungen einzubeziehen und damit zusätzliche N Grenzkostenbedingungen eines Produktionsprogrammes zur Spezifikation zur Verfügung zu stellen. HECKELEI & BRITZ verwenden diese grundsätzliche Idee zur Schätzung regionaler Kostenfunktionen auf Basis von Querschnittsdaten.

Zunächst wollen wir jedoch argumentieren, dass die PMP-Prozedur nicht zur optimalen Ausnutzung der zusätzlichen Dateninformation geeignet ist, da die aus dem ersten Schritt abgeleiteten Grenzkostenbedingungen keine konsistente Schätzung der Parameter erlauben. Zur Verdeutlichung ist es notwendig, sich PMP aus der Sicht eines Ökonometrikers anzuschauen: Dazu gehört die Vorstellung eines „wahren“ Modells, oder zumindest die Annahme, das ein

bestimmtes Modell den wahren Datengenerierungsprozess ausreichend gut annähern kann. Die Vielzahl der PMP-Nutzer glaubte dies anscheinend in Bezug auf das resultierende nicht-lineare Modell, mit dem anschließend ökonomische Analysen durchgeführt wurden. Nehmen wir beispielsweise das resultierende quadratische Modell (3) so lässt sich zeigen, dass unter der Annahme ausschließlich positiver Aktivitätsumfänge und bindender Ressourcenbeschränkungen im Optimum die Schattenpreise den Wert $\lambda = (\mathbf{A}\mathbf{Q}^{-1}\mathbf{A}')^{-1}(\mathbf{A}\mathbf{Q}^{-1}(\mathbf{p}-\mathbf{d})-\mathbf{b})$ annehmen.

Im ersten Schritt von PMP ergibt sich allerdings etwas anderes: Zur Illustration partitionieren wir \mathbf{I} in zwei Untervektoren, den $((N-M)\times 1)$ Vektor der "vorzüglichen" Aktivitäten, \mathbf{I}^v , welche gegen die Kalibrierungsbeschränkungen laufen, und ein $(M\times 1)$ Vektor "marginaler" Aktivitäten, \mathbf{I}^m , deren Umfänge durch die Ressourcenbeschränkung gehalten werden. Aus den Kuhn-Tucker Bedingungen können nun die Dualwertgleichungen

$$(4) \quad (a) \lambda = (\mathbf{A}^m)'^{-1}(\mathbf{p}^m - \mathbf{c}^m), \quad (b) \rho^v = \mathbf{p}^v - \mathbf{c}^v - \mathbf{A}^v \lambda, \quad (c) \rho^m = \mathbf{0},$$

hergeleitet werden. Wir stellen fest, dass sich λ ausschließlich über die Zielfunktionsbeiträge der marginalen Aktivitäten definieren lässt und sich somit generell von den Werten des quadratischen Modells unterscheiden. Da in "Schritt 1" ρ simultan mit λ bestimmt wird (siehe (4)) und zur Bestimmung der \mathbf{GK}^v verwendet wird (siehe (2)), sind auch Letztere i.A. verzerrt. Schließlich werden im "Schritt 2" diese verzerrten \mathbf{GK}^v zur Bestimmung der Parameter herangezogen, was eine konsistente Schätzung der Parameter einer nicht-linearen Zielfunktion mit dem PMP-Ansatz i.d.R. unmöglich macht.²

Die hier vorgeschlagene „Allgemeine Alternative“ zu PMP beruht auf einem simplen Prinzip, nämlich der direkten Verwendung der Optimalitätsbedingungen des gewünschten Programmierungsmodells. Es ist kein „Schritt 1“ zur Bestimmung der Dualwerte von Kalibrierungsbeschränkungen notwendig. Stattdessen kann die methodische Inkonsistenz durch eine simultane Schätzung von Schattenpreisen und Modellparametern vermieden werden. Wie sich an einem unten präsentierten Beispiel noch zeigt, ist diese Vorgehensweise nicht nur dazu geeignet Parameter von typischen Programmierungsmodellen zu schätzen, sondern erlaubt auch eine alternative Schätzung von dualitätstheoretisch basierten Verhaltensfunktionen. In diesem Zusammenhang besteht der Unterschied zwischen Programmierungs- und ökonometrischen Modellen dann nur noch in der bei der Simulation verwendeten Modellform. Es überrascht vielleicht nicht, dass die innovativsten PMP-Vertreter PARIS and HOWITT (2001) dieses Prinzip zur *Kalibrierung* eines Angebotsmodells bereits (unbewusst) genutzt haben: Ihr “Symmetric Positive Equilibrium Problem” (SPEP) kalibriert ein Multi-Input-Multi-Output Gewinnmaximierungsproblem auf Basis von Grenzkostenbedingungen. Allerdings hatten die Autoren nicht erkannt, dass die von Ihnen benutzte erste Phase zur Bestimmung der Dualwerte von Kalibrierungsbeschränkungen vollkommen irrelevant für die Vorgehensweise war.³

Im nächsten Abschnitt soll das grundsätzliche Prinzip der Schätzung von Programmierungsmodellen anhand dreier Gewinnmaximierungsmodelle erläutert werden. Monte Carlo Simulationen auf Basis von GME-Schätzungen geben Hinweise auf die Konsistenz der Vorgehensweise und zeigen zusätzlich den Einfluss von Vorinformation auf die Qualität der Schätzungen. Die zu schätzenden Programmierungsmodelle stellen nicht notwendigerweise besonders geeignete Modelle der Agrarangebotsanalyse aus unserer Sicht dar, sondern sind mit dem Wunsch ausgewählt, einen Bogen zwischen der Literatur zu Programmierungsmodellen und der zu ökonometrisch geschätzten Verhaltensfunktionen zu spannen.

² Siehe auch WILD (2000), der dies beispielhaft illustriert.

³ Für Interessierte: Die linken Seiten der Datenbedingungen, die Dualwerte von Kalibrierungsbeschränkungen aus Phase 1 und sogenannte "limiting inputs" enthalten und zur ME-Bestimmung der Kostenfunktion genutzt werden (siehe Gleichungen 26 und 27, S.93), können durch die aus den Daten bekannten Produktpreise und Gesamtinputmengen aufgrund der Eigenschaften der Phase 1 ersetzt werden.

3 Beispiele zur Schätzung von Programmierungsmodellen

3.1 Modell mit quadratischer Kostenfunktion

Dieser Unterabschnitt beschäftigt sich mit der Parameterschätzung für das im PMP Zusammenhang oft verwendete und oben schon beschriebene Optimierungsmodell mit einer quadratischen Kostenfunktion. Für das quadratische Programmierungsmodell („QP-Modell“) wollen wir aus Vereinfachungsgründen nur die Menge des verfügbaren Landes, als fix ansehen, so dass der Vektor der Ressourcenbeschränkungen und der korrespondierenden Schattenpreise in (3) zum Skalar werden. Ausserdem ersetzen wir in den Vektor der Preise \mathbf{p} durch einen Vektor von Deckungsbeiträgen \mathbf{db}^4 und erhalten somit

$$(5) \quad \begin{aligned} \max_{\mathbf{l}} Z &= \mathbf{db}'\mathbf{l} - \mathbf{d}'\mathbf{l} - 0,5\mathbf{l}'\mathbf{Q}\mathbf{l} \quad \text{so dass} \\ \mathbf{u}'_N\mathbf{l} &\leq b \quad [\lambda], \quad \mathbf{l} \geq \mathbf{0} \end{aligned}$$

wobei \mathbf{u}_N einen $(N \times 1)$ Summationsvektor darstellt.

Als Schätzmethode wird hier die „Verallgemeinerte Maximum Entropie“ („Generalised Maximum Entropy“, GME) Methode nach Golan et al. (1996) angewandt, die eine flexible Einbeziehung von Vorinformationen erlaubt.⁵ Bei Gültigkeit von Modell (5) erfüllt die beobachtete Landallokation die Bedingungen erster Ordnung für jeden Datenpunkt $t = 1, \dots, T$. Geht man von einem additiven stochastischen Messfehler der endogenen Variablen aus und bezeichnet den $(N \times 1)$ Vektor der Fehlerterme für die Beobachtung t als \mathbf{e}_t , so lassen sich diese Optimalitätsbedingungen schreiben als⁶

$$(6) \quad \begin{aligned} \mathbf{db}_t^\circ - \lambda_t \mathbf{u}_N - \mathbf{d} - \mathbf{Q}(\mathbf{l}_t^\circ - \mathbf{e}_t) &= 0 \\ \mathbf{u}'(\mathbf{l}_t^\circ - \mathbf{e}_t) &= b_t \end{aligned}$$

Zur Anwendung der GME Methode reparametrisieren wir die Fehlervektoren als Erwartungswert einer diskreten Wahrscheinlichkeitsverteilung. Die „Stützpunktmatrix“ \mathbf{V} begrenzt in unserem Fall den Wertebereich der Fehlerterme auf ± 10 Standardabweichungen der Messfehler, σ_i , um den Nullpunkt.⁷ Im Fall des unten verwendeten Simulationsexperiments mit 3 Anbaukulturen können die drei Fehlerterme dann durch die Multiplikation von \mathbf{V} mit einem $((N \cdot 2) \times 1)$ -Vektor von Wahrscheinlichkeiten \mathbf{w}_t dargestellt werden als

⁴ Somit geht dieses Modell davon aus, dass die quadratische Kostenfunktion „irgendwelche“ nicht linearen Kosten widerspiegelt, die nichts mit den variablen Inputs pro Flächeneinheit zu tun haben. Diese mangelnde Rationalisierung des Modells ist analog zu vielen PMP Anwendungen. Mit der Verwendung soll hier aber auf keinen Fall eine Aussage über die Vorzüglichkeit dieser theoretischen Modellformulierung bei der Agrarangebotsanalyse gemacht werden.

⁵ Es soll aber darauf hingewiesen werden, dass in dem hier vorgestellten Zusammenhang überbestimmter Gleichungssysteme auch andere Schätzmethode als GME zum Einsatz kommen können.

⁶ Hier und nachfolgend gilt die Konvention, dass Gleichungen ohne formellen Hinweis für alle in den jeweiligen Indizes enthaltenen Elemente definiert sind. Das heißt in diesem Fall für alle $t = 1, \dots, T$.

⁷ Die Frage der angemessenen Weite des Stützpunktbereichs der Fehlerterme ist eine viel diskutierte und nicht endgültig geklärt Frage. GOLAN et al. schlagen die „3-Sigma“ Regel vor. PRECKEL (2001) plädiert für einen möglichst großen Bereich zur Annäherung an das Schätzverhalten des Kleinstquadrateschätzers. Da in diesem Artikel die allgemeine Vorgehensweise bei der Schätzung und nicht die Eigenschaften des Entropieschätzers im Vordergrund stehen, geben wir uns mit einer „ad hoc“ 10-Sigma Spezifikation zufrieden

$$(7) \quad \mathbf{e}_t = \mathbf{V}\mathbf{w}_t = \begin{bmatrix} -10\sigma_1 & 10\sigma_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -10\sigma_2 & 10\sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & -10\sigma_3 & 10\sigma_3 \end{bmatrix} \begin{bmatrix} w_{11t} \\ w_{12t} \\ w_{21t} \\ w_{22t} \\ w_{31t} \\ w_{32t} \end{bmatrix}$$

Zusätzlich zu den Restriktionen (6) und (7) ergibt sich die vollständige GME Schätzprozedur von Parametern und Schattenpreisen als⁸

$$(8) \quad \max_{\mathbf{w}_t, \mathbf{e}_t, \mathbf{Q}, \mathbf{L}, \lambda_t} H(\mathbf{w}_t) = -\sum_{t=1}^T \mathbf{w}_t' \ln \mathbf{w}_t$$

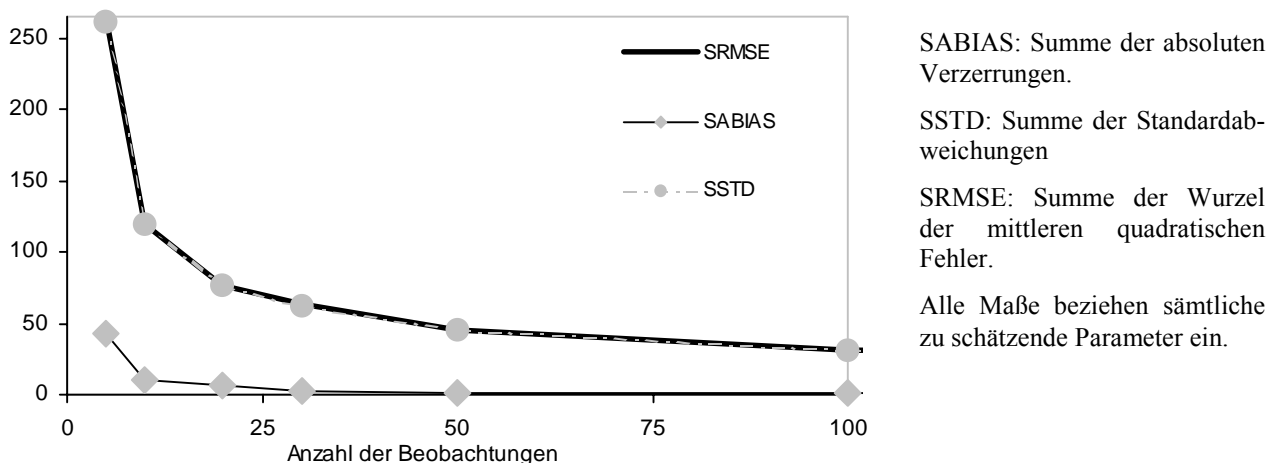
$$(9) \quad \mathbf{Q} = \mathbf{L}\mathbf{L}' \text{ mit } L_{ij} = 0 \quad \forall j > i$$

$$(10) \quad \sum_{k=1}^2 w_{kit} = 1$$

wobei $H(\mathbf{w}_t)$ die Entropie ist, in (9) die positive Definitheit von \mathbf{Q} mit Hilfe einer Cholesky Faktorisierung gewährleistet wird und Gleichung (10) die Addition der Wahrscheinlichkeiten auf 1 erzwingt.

Folgendes Simulationsexperiment wird zur Überprüfung der Schätzqualität durchgeführt:⁹ Auf Basis der Produkt- und Inputdifferenzierung in HOWITT (1995b) wird ein Datensatz mit T Beobachtungen durch Optimierung des Modells (5) für T verschiedene Vektoren $\mathbf{d}\mathbf{b}_t$ und gegebene Parameter \mathbf{Q} generiert. Das verfügbare Land wird über die Beobachtungen hinweg konstant gehalten. Zu den optimalen Landallokationen l_{1t}^* und l_{2t}^* der ersten beiden Früchte werden dann normalverteilte Meßfehler mit einer Standardabweichung von 2% der durchschnittlichen Landallokation addiert, so dass sich die „beobachteten“ Landallokationen als $l_{1t}^o = l_{1t}^* + e_{1t}$ und $l_{2t}^o = l_{2t}^* + e_{2t}$ ergeben. Um sicher zu stellen, dass die Landrestriktion bei den beobachteten Produktionsumfängen bindend ist, lassen wir $l_{3t}^o = b_t - l_{1t}^* - l_{2t}^*$.

Grafik 1 QP-Modell - Monte Carlo Simulationsergebnisse ohne a-priori Information



Quelle: Eigene Berechnungen.

⁸ Für den vorliegenden Fall nur einer Ressourcenbeschränkung ist der Vektor \mathbf{d} nicht identifiziert, weshalb dessen Elemente im nachfolgenden Simulationsexperiment bei der Datengeneration und Schätzung auf „0“ gesetzt werden.

⁹ Viele Details der Vorgehensweise können hier aus Platzgründen nicht dargestellt werden. Dies gilt noch mehr für die nachfolgenden Beispielmmodelle. Eine Beschreibung mit dem Anspruch der „Reproduzierbarkeit“ findet sich in HECKELEI (2001).

Für jeden generierten Datensatz werden die Modellparameter mit dem Entropieverfahren geschätzt und bei einer ausreichenden Anzahl von Wiederholungen die Güte der Schätzung durch bekannte Maße wie die absolute Verzerrung („ABIAS“ = absoluter Wert der Differenz zwischen mittlerem Schätzwert und dem wahren Wert des Parameters), der Standardabweichung („STD“) und der Wurzel des Mittlerem Quadratischen Fehlers („RMSE“) berechnet. Wegen der Vielzahl der Parameter findet i.d.R. eine Summierung über alle Parameter statt.

Grafik 1 präsentiert die Monte Carlo Simulationsergebnisse im Hinblick auf die Elemente der Matrix Q für unterschiedliche Stichprobengrößen. Die Summe aller RMSE sinkt mit dem Anstieg der Stichprobengröße deutlich, was auf eine mögliche Konsistenz des Schätzverfahrens hindeutet. Die Verzerrung der Schätzer ist schon bei einer Stichprobengröße von 20 zu vernachlässigen (SABIAS). Insgesamt ist die Varianz der Schätzer – hier gemessen mit der Summe der Standardabweichungen (SSTD) – für den mittleren Fehler von wesentlich größerer Bedeutung.

Eine empirisch relevante Möglichkeit der Einbeziehung von a-priori Informationen bei der Schätzung ist die Maßgabe der Größenordnung von Elastizitäten. Dabei kann man häufig auf verschiedene Studien, die eine gewisse Vergleichbarkeit hinsichtlich der empirischen Zusammenhänge erlauben, zurückgreifen. Unten wird jedoch deutlich, dass die Präzision der Vorinformation durchaus begrenzt sein kann, um trotzdem die Qualität der Schätzergebnisse bei kleinen Stichprobengrößen zu verbessern. Die technische Umsetzung erfolgt prinzipiell über Reparametrisierung der Elastizitäten durch Stützpunkte und Wahrscheinlichkeiten entsprechend der oben dargestellten Vorgehensweise bei den Fehlertermen. Das Entropie-Kriterium sorgt, im Rahmen des durch die Nebenbedingungen definierten Lösungsraumes, tendenziell für eine Schätzgröße, die möglichst nah an den a-priori Werten (Erwartungswerte bei Gleichverteilung der Wahrscheinlichkeiten) liegen. Je enger der Stützpunktbereich, desto stärker werden Abweichungen vom Erwartungswert bestraft. Die Variation des Stützpunktbereichs erlaubt somit eine Anpassung der Präzision bzw. des Einflusses der a-priori Information. Der Ansatz ist analog zur üblichen Vorgehensweise im Rahmen von GME, bei der allerdings in der Theorie (GOLAN et al.) und agrarökonomischen Anwendungen (z.B. LENCE & MILLER 1998a und 1998b; OUDE LANSINK 1999; ZHANG & FAN 2001) bisher nur direkte Restriktionen des Parameterraums verwendet wurden. Die hier genutzten Beschränkungen bestimmter Funktionen von Parametern liegen allerdings oft näher an typischerweise verfügbaren Vorinformationen.

Nach diesen technischen Anmerkungen nun zur Formulierung der a-priori Information im Rahmen der Simulationsexperimente. Der Stützpunktbereich der Elastizitäten wird mit einer Spanne von 4 so definiert, dass eine substantielle Variationsbreite der Schätzergebnisse ohne starke Bestrafung durch die Zielfunktion erlaubt ist. In Bezug auf die a-priori Erwartungswerte der Elastizitäten werden zwei verschiedene Versionen simuliert: zum Einen mit a-priori Erwartungswerten, die dem wahren Wert im Mittelpunkt der Daten entsprechen, zum Anderen mit solchen die um 0,3 nach oben verschoben sind.

Tabelle 1 QP-Modell - RMSE der geschätzten Deckungsbeitragselastizitäten der Landallokation bei unterschiedlichen a-priori Informationen

| A-priori Information | Anzahl der Beobachtungen (T) | | | | |
|----------------------|------------------------------|-------|-------|-------|-------|
| | 5 | 10 | 20 | 30 | 50 |
| "ohne" | 0.187 | 0.110 | 0.071 | 0.055 | 0.045 |
| "wahr" | 0.158 | 0.105 | 0.063 | 0.055 | 0.045 |
| "falsch" | 0.163 | 0.105 | 0.063 | 0.055 | 0.045 |

Quelle: Eigene Berechnungen. Der Wert der wahren Deckungsbeitragselastizität ist 1.03.

Tabelle 1 berichtet stellvertretend die RMSE der Deckungsbeitragselastizität eines Produktes bei verschiedenen Stichprobengrößen. Daran wird zunächst deutlich, dass den oben sichtbaren deutlichen Schwankungen der geschätzten Parameter relativ stabile Elastizitäten auch schon bei einer kleinen Anzahl von Beobachtungen gegenüberstehen. Dennoch erkennt man die

potentiellen Vorteile einer Einbeziehung von Vorinformationen: Der mittlere Schätzfehler nimmt bei kleinen Stichproben ab, und zwar für beide Formulierungen der a-priori Information. Die Reduktion der Schätzvarianz gleicht offensichtlich die eingeführte Verzerrung bei der „falschen“ Vorinformation mehr als aus. Ab einer Stichprobengröße von 30 sind keine Unterschiede mehr zu erkennen und alle 3 Varianten laufen mit Zunahme der Zahl der Beobachtungen auf die wahren Werte zu.

3.2 Inputallokation auf Basis von Produktionsfunktionen

In diesem Abschnitt wollen wir ein Programmierungsmodell betrachten, das unter Verwendung einer funktionalen Abbildung der Produktionstechnologie eine simultane Allokation variabler Inputs und mehrerer fixer Ressourcen auf die einzelnen Produktionsprozesse vornimmt. Die allgemeine Form des gewünschten Gewinnmaximierungsmodells ist gegeben durch

$$(11) \quad \max_{\mathbf{x}, \mathbf{b}} Z = \sum_{i=1}^N p_i f_i(x_{ik}, b_{ij}) - \sum_{i=1}^N \sum_{k=1}^K q_k x_{ik} \quad \text{so dass} \quad \sum_{i=1}^N b_{ij} = b_j \quad [\lambda_j]$$

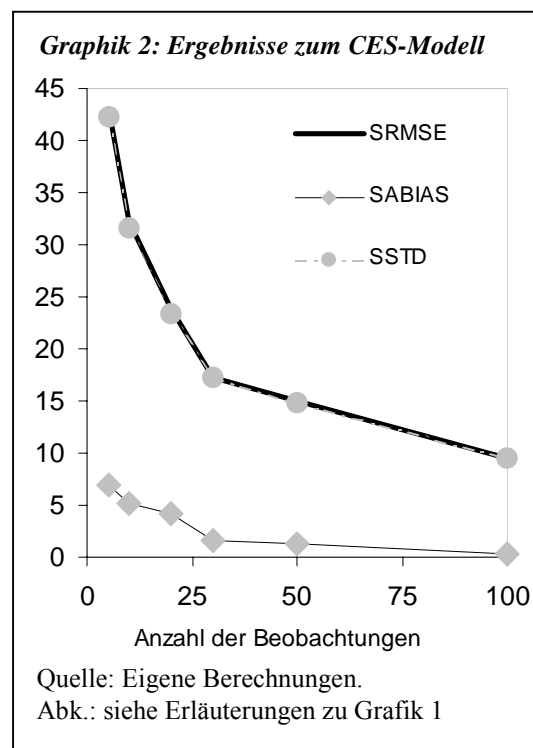
wobei i, j, k jeweils die Indizes der Aktivitäten, fixen sowie variablen Inputs sind, q_k und x_{ik} Preise und alloziierte Mengen der variablen Inputs und b_{ij} und b_j die alloziierten und gesamt verfügbaren Mengen der fixen Faktoren darstellen. Die Transformation der Inputmengen in Output y_i ist durch

$$(12) \quad y_i = f_i(x_{ik}, b_{ij})$$

gegeben. Die Optimalitätsbedingungen erster Ordnung setzen sich zusammen aus den Ressourcenbeschränkungen, den Wertgrenzproduktivitätsbedingungen variabler Faktoren und den Schattenpreisgleichungen der fixen Faktoren:

$$(13) \quad \sum_{i=1}^N b_{ij} = b_j, \quad \frac{\partial Z}{\partial x_{ik}} = p_i \frac{\partial f_i}{\partial x_{ik}} - q_k = 0, \quad \frac{\partial Z}{\partial b_{ij}} = p_i \frac{\partial f_i}{\partial b_{ij}} - \lambda_j = 0.$$

Als Bsp. wird für die Produktionsfunktion die „Constant Elasticity of Substitution“ (CES) Funktion gewählt, wobei zwischen zwei variablen Inputs sowie den fixen Faktoren Land und Wasser unterschieden wird. Dieses System nach den Inputnachfrage- und Angebotsfunktionen analytisch abzuleiten, dürfte sehr aufwendig, wenn nicht gar unmöglich sein. Stattdessen können die Gleichungen (12) bis (13) als Datenbeschränkungen zur Schätzung der Parameter des Programmierungsmodells (11) verwendet werden, wodurch sich ein erheblicher Flexibilitätsvorteil hinsichtlich der Wahl der Funktionsform und der Modellstruktur ergibt. Der Modellaufbau ist analog zu dem PMP-CES Ansatz von HOWITT (1995b). Allerdings enthält die vorliegende Version keine zusätzlichen nicht-linearen Kostenterme in der Zielfunktion (und der Schätzansatz bedarf natürlich auch keiner Bestimmung der Dualwerte von Kalibrierungsbeschränkungen aus der ersten Phase von PMP). Ein weiterer Unterschied ist die Formulierung



sinkender Skalenerträge, damit positive Produktionsumfänge aller Früchte möglich sind.¹⁰

Für unsere Monte-Carlo Simulation werden im Datengenerierungsprozess die „beobachteten“ Outputmengen und Inputallokationen mit normalverteilten Meßfehlern versehen (jeweils mit 10% und 2% Standardabweichung vom Mittelwert), wobei wiederum die „falsch gemessenen“ allozierten Mengen der fixen Inputs sich zu den beobachteten Aggregaten exakt aufaddieren. Ein entsprechender GME Schätzansatz führt durch Stützpunkte und Wahrscheinlichkeiten parametrisierte Fehlerterme in die Gleichungen ein.¹¹ Das Entropie-Kriterium bestimmt dann, unter Einhaltung der Datenbeschränkungen, die Parameter der Produktionsfunktionen und λ . Grafik 2 zeigt die absolute Verzerrung, Standardabweichung und den mittleren quadratischen Fehler jeweils als Summe über alle Parameter der Produktionsfunktionen. Die im Vergleich zum SABIAS hohe Varianz könnte darauf hindeuten, dass die Schätzgüte durch Einführung von plausibler a-priori Information verbessert werden kann.

3.3 Allokation Fixer Faktoren auf Basis von Gewinnfunktionen

Das letzte Beispiel zur Schätzung eines Programmierungsmodells behält die grundsätzliche Modellstruktur hinsichtlich Verhaltens- und Technologieannahme des vorherigen Unterabschnitts bei, bedient sich aber bei der Bestimmung von Output- und variablen Inputmengen den Konzepten der Dualitätstheorie. Betrachtet man nur einen fixen Faktor, so ist die Spezifikation analog zu GUYOMARD et al. (1996) und MORO & SCHOKAI (1999) die entsprechende ökonomische Schätzungen von Systemen mit Angebots- und expliziten Landallokationsfunktionen vorgestellt haben. Hier soll zum Einen auf die vollkommene Äquivalenz unserer Vorgehensweise im Hinblick auf die Schätzergebnisse hingewiesen, zum Andern die Vorteile durch Flexibilisierung bei der Funktionswahl und komplexeren Modellstrukturen verdeutlicht werden. Das gewünschte Programmierungsmodell hat die folgende Struktur:

$$(14) \quad \max_{\mathbf{l}} Z = \sum_{i=1}^N \pi_i(\mathbf{p}_i, \mathbf{q}, l_i) \quad \text{so dass} \quad \sum_{i=1}^N l_i = b \quad [\lambda]$$

wobei

$$(15) \quad \pi_i(\mathbf{p}_i, \mathbf{q}, l_i) = \max_{y_i, \mathbf{x}_i} \left[p_i y_i - \sum_{k=1}^K q_k x_{ik} \quad \text{so dass} \quad y_i = f_i(x_{ik}, l_i) \right]$$

eine restringierte Gewinnfunktion der Produktion der Frucht i bei gegebener Landallokation l_i darstellt. Das Modell (14) verteilt also das verfügbare Land b auf die einzelnen Produktionsverfahren zur Maximierung des Gesamtgewinns, wobei der Gewinn in den einzelnen Fruchtarten durch spezifische Gewinnfunktionen bestimmt wird. Dementsprechend ist die optimale Landallokation gegeben, wenn die Grenzgewinne des Landes in allen Verwendungen gleich sind, also die Bedingungen erster Ordnung

$$(16) \quad \frac{\partial \pi_i(\mathbf{p}_i, \mathbf{q}, l_i)}{\partial l_i} - \lambda = 0$$

gelten. Bei bestimmten Funktionsformen von $\pi(\cdot)$ ist eine Auflösung des Systems (16), unter zusätzlicher Verwendung der Landbeschränkung, nach expliziten Landallokationsfunktionen möglich, die die Landnachfrage in Abhängigkeit exogener Modellgrößen bestimmen. GUYOMARD et al. beschreiben diese Herleitung und schätzen ein System von

¹⁰ Eine Verwendung konstanter Skalenerträge hätte hier eine Spezialisierung zur Folge, da der in jeder Frucht maximal zu erzielende Gewinn/Flächeneinheit unabhängig von der Flächenallokation ist. Dadurch kann die Anzahl positiver Produktionsumfänge die Anzahl der fixen Faktoren im Optimum wie beim LP nicht überschreiten.

¹¹ Die Einführung von Fehlertermen um die endogenen Variablen \mathbf{x}_i und \mathbf{b}_j , erlauben eine mit dem ökonomischen Modell konsistente simultane Schätzung der Inputallokation. Der vermuteten Qualität „beobachteter“ Inputallokationen kann dann durch Anpassung der Größe des Stützbereichs für die Fehlerterme Rechnung getragen werden.

Landallokationsgleichungen und durch Hotelling's Lemma bestimmte Angebotsgleichungen der Form

$$(17) \quad \frac{\partial \pi_i(\mathbf{p}_i, \mathbf{q}, l_i)}{\partial p_i} = y_i(\mathbf{p}_i, \mathbf{q}, l_i)$$

basierend auf normalisiert quadratischen Gewinnfunktionen. Zwar ist dieses System linear, aber die Regressionskoeffizienten unterliegen nicht-linearen Restriktionen über Gleichungen hinweg. Bei unserer Vorgehensweise findet die Auflösung nach den Landallokationsgleichungen nicht statt. Statt dessen werden direkt die Optimalitätsbedingungen (16) in Verbindung mit (17) als Datenrestriktionen eines GME Ansatzes analog zu den beiden vorigen Abschnitten genutzt. Solange das statistische Modell und das verwendete Schätzkriterium gleich ist, müssen bei diesem Ansatz aufgrund der mathematische Äquivalenz der Datenrestriktion genau die gleichen Schätzwerte der Parameter herauskommen wie bei der Vorgehensweise von GUYOMARD et al., was von uns auch auf Basis des GME-Ansatzes und eines nicht-linearen Kleinstquadrat-schätzers nachgewiesen wurde.

Welchen Vorteil bringt dann die Schätzung auf Basis der Optimalitätsbedingungen und die anschließende Verwendung eines Programmierungsmodells? Die folgenden Punkte sind zu nennen: 1) Die Flexibilität bei der Wahl der Funktionsform für $\pi_i(\mathbf{p}_i, \mathbf{q}, l_i)$ wird wesentlich erhöht, da es keiner geschlossenen Lösung der Landallokationsfunktionen bedarf; 2) Aus dem gleichen Grund ist eine komplexere Modellstruktur mit mehreren fixen Faktoren wie in Abschnitt 3.2 oder anderen im Datenzeitraum relevanten Nebenbedingungen kein grundsätzliches Hindernis für die ökonometrische Schätzung der Parameter mehr; 3) Die Formulierung des resultierenden Simulationsmodells als explizites Optimierungsmodell erlaubt die flexible Einbeziehung von zusätzlichen relevanten Nebenbedingungen der Allokation im Simulationszeitraum ohne dass die strukturelle Gültigkeit der Schätzergebnisse notwendigerweise aufgehoben wird.

Zum Abschluß sollen auch hier die Eigenschaften eines entsprechenden GME Schätzers mit einem Simulationsexperiment überprüft werden. Dazu wurden für verschiedene Stichproben-größen optimale Angebotsmengen und Landallokationen mit vorigen Modellen entsprechenden Messfehlern versehen und anschließend eine Parameterschätzung ohne Verwendung von a-priori Informationen durchgeführt. Das Ergebnis ist in Tabelle 3 wiedergegeben. Der Verlauf der über alle Parameter summierten Maße bei Zunahme der Beobachtungen deutet erneut auf ein konsistentes Verhalten des GME-Schätzers hin. Die hohe Varianz bei kleinen Stichprobengrößen läßt auch hier eine Qualitätsverbesserung der Schätzgüte durch Einführung von realistischen a-priori Informationen – z.B. in Form von Elastizitäten – möglich erscheinen.

Tabelle 2 Gewinnfunktionsmodell – Monte Carlo Ergebnisse ohne a-priori Information

| Gütemaße | Beobachtungen | | | | | | |
|----------|---------------|------|------|-----|-----|-----|-----|
| | 4 | 5 | 10 | 20 | 30 | 50 | 100 |
| SRMSE | 2965 | 2888 | 1212 | 570 | 462 | 346 | 253 |
| SABIAS | 914 | 900 | 417 | 222 | 159 | 102 | 57 |
| SSTD | 2715 | 2672 | 1102 | 516 | 426 | 325 | 242 |

Quelle: Eigene Berechnungen

4 Schlussfolgerungen

Der vorliegende Beitrag stellt ein Prinzip zur Schätzung von Programmierungsmodellen auf Basis der Optimalitätsbedingungen vor. Die Anwendung wird an drei verschiedenen Beispielen demonstriert und die grundsätzliche Funktionsfähigkeit mit Monte Carlo Simulationsergebnissen nachgewiesen. Die Methode erlaubt gleichzeitig die Schätzung komplexerer Modelle mit flexiblerer funktionaler Spezifikation als bei bisherigen ökonometrischen Schätzungen dualitätstheoretisch fundierter Verhaltensfunktionen. Der verwendete GME Schätzansatz zeigt

die möglichen Vorteile einer Einbeziehung von a-priori Informationen bei kleinen Stichprobengrößen durch Reduktion der Stichprobenvarianz. Systematischere Untersuchungen zur Formulierung von a-priori Informationen und den Konsequenzen ihrer Verwendung für die Schätzqualität wären wünschenswert. Von großem Interesse aus Sicht der Autoren ist auch die grundsätzliche Möglichkeit der Schätzung über Beobachtungen mit bindenden und nicht-bindenden Nebenbedingungen hinweg. Dies muss aber im Hinblick auf die Praktikabilität noch nachgewiesen werden.

5 Literatur

- CYPRIS, Ch. (2000): Positiv Mathematische Programmierung (PMP) im Agrarsektormodell RAUMIS. Dissertation, Universität Bonn.
- GOLAN, A., G. JUDGE, und D. MILLER (1996): *Maximum Entropy Econometrics*, Chichester UK: Wiley.
- GUYOMARD, H., M. BAUDRY und A. CARPENTIER (1996): Estimating crop supply response in the presence of farm programmes: application to the CAP. *European Review of Agricultural Economics*, 23: 401-420.
- HECKELEI, T. (2001): The Calibration and Estimation of Programming Models in Agricultural Supply Analysis, in Vorbereitung.
- HECKELEI, T, und W. BRITZ (2001): Positive Mathematical Programming with Multiple Data Points: A Cross-Sectional Estimation Procedure. *Cahiers d'Economie et Sociologie Rurales*, Herbst 2001, im Druck.
- HELMING, J. F. M., L. PEETERS, and P.J.J. VEENDENDAAL (2001): Assessing the Consequences of Environmental Policy Scenarios in Flemish Agriculture. In: Heckeley, T., H.P. Witzke, and W. Henrichsmeyer (Eds.): *Agricultural Sector Modelling and Policy Information Systems*. Proceedings of the 65th EAAE Seminar, March 29-31, 2000 at Bonn University, Vauk Verlag Kiel, pp. 237-245.
- HOWITT, R.E. (1995a): Positive Mathematical Programming. *American Journal of Agricultural Economics*, 77(2): 329-342.
- HOWITT, R. E. (1995b): A Calibration Method for Agricultural Economic Production Models. In: *Journal of Agricultural Economics* 46: 147-159.
- LENCE, H.L. und D. MILLER (1998a): Estimation of Multioutput Production Functions with Incomplete Data: A Generalized Maximum Entropy Approach. *European Review of Agricultural Economics*, 25:188-209.
- LENCE, H.L. und D. MILLER (1998b): Recovering Output Specific Inputs from Aggregate Input Data: A Generalized Cross-Entropy Approach. *American Journal of Agricultural Economics*, 80(4):852-867.
- OUDE LANSINK, A. (1999): Generalized Maximum Entropy and heterogeneous technologies. *European Review of Agricultural Economics*, 26:101-115.
- PARIS, Q., and R.E. HOWITT (1998): An Analysis of Ill-Posed Production Problems Using Maximum Entropy. *American Journal of Agricultural Economics*, 80(1):124-138.
- PARIS, Q., and R.E. HOWITT (2001): The Multi-Input Multi-Output Symmetric Positive Equilibrium Problem. In: Heckeley, T., H.P. Witzke, and W. Henrichsmeyer (Eds.): *Agricultural Sector Modelling and Policy Information Systems*. Proceedings of the 65th EAAE Seminar, March 29-31, 2000 at Bonn University, Vauk Verlag Kiel, pp. 88-100.
- WILD, L. (2000): Schätzung von Kostenfunktionen im Rahmen der Positiv Mathematischen Programmierung. Diplomarbeit an der Universität Bonn am Lehrstuhl für Volkswirtschaftslehre, Agrarpolitik und Landwirtschaftliches Informationswesen.
- ZHANG, X. und S. FAN (2001): Estimating Crop-Specific Production Technologies in Chinese Agriculture: A Generalized Maximum Entropy Approach. *American Journal of Agricultural Economics*, 83(2): 378-388.