

# Analyse der Repräsentativität im schweizerischen landwirtschaftlichen Buchhaltungsnetz

Messung und Verbesserung der Schätzqualität  
ökonomischer Kennzahlen in der Zentralen  
Auswertung von Buchhaltungsdaten

**Doctoral Thesis**

**Author(s):**

Meier, Beat

**Publication date:**

2005

**Permanent link:**

<https://doi.org/10.3929/ethz-a-004943484>

**Rights / license:**

In Copyright - Non-Commercial Use Permitted

Diss ETH No. 15868

# **Analyse der Repräsentativität im schweizerischen landwirtschaftlichen Buchhaltungsnetz**

Messung und Verbesserung der Schätzqualität ökonomischer  
Kennzahlen in der Zentralen Auswertung von Buchhaltungsdaten

ABHANDLUNG  
zur Erlangung des Titels  
DOKTOR DER WISSENSCHAFTEN  
der  
EIDGENÖSSISCHEN TECHNISCHEN HOCHSCHULE ZÜRICH

vorgelegt von

BEAT MEIER  
Dipl Ing.-Agr. ETH  
geboren am 1.12.1964

von  
Hohenrain LU

Angenommen auf Antrag von

Prof. Dr. Bernard Lehmann, Referent  
Prof. Dr. Walter Meier, Korreferent

Zürich 2005

## Zusammenfassung

Die Zentrale Auswertung landwirtschaftlicher Buchhaltungsdaten in der Schweiz beruht bis 1998 auf einer Stichprobe sogenannter Testbetriebe. Diese Stichprobe stellt eine bewusste Selektion betriebswirtschaftlich überdurchschnittlicher Betriebe dar. Der 1999 neu formulierte gesetzliche Auftrag enthält demgegenüber die Zielsetzung, die wirtschaftliche Situation in der Landwirtschaft mit einer Stichprobe repräsentativer Referenzbetriebe abzubilden. Ausgehend vom vorhandenen Datenmaterial wird dieses Ziel mit dem Massnahmenpaket „Methode 2000“ angestrebt. Die wichtigsten methodischen Änderungen betreffen die Abgrenzung der Grundgesamtheit, die Entwicklung einer neuen Betriebstypologie, die Einführung eines geschichteten Auswahlplanes und einer Gewichtung der einzelbetrieblichen Ergebnisse.

Aufgrund der fehlenden Zufallsauswahl der Referenzbetriebe sind statistische Schlüsse von der Stichprobe auf die Grundgesamtheit problematisch. Für die heute geschätzten Mittelwerte muss mit erheblichen Verzerrungen gerechnet werden. Die Überprüfung der Qualität dieser Schätzungen erfolgt durch mehrere, methodisch unterschiedliche Ansätze. Die Ergebnisse zeigen, dass durch die mangelhafte Vertretung der Spezialkultur- und der Veredlungsbetriebe sowie durch den Ausschluss der Betriebsgemeinschaften von den Referenzbetrieben, die heute geschätzten landwirtschaftlichen Einkommen gegenüber der Grundgesamtheit eine negative Verzerrung in der Grössenordnung von 2% aufweisen, d.h. die aktuelle Schätzung aus diesem Grunde zu tief ausfällt. Ein Kalibrierungsmodell zur Korrektur verzerrter Hilfsvariablen (Flächen, Alter, Bildung, Bio-Betriebe) lässt demgegenüber auf eine positive Verzerrung des landwirtschaftlichen Einkommens von 5% bis 6% schliessen. Die in der aktuellen Stichprobe überdurchschnittliche Ausbildung der betriebsleitenden Personen stellt dabei eine Hauptursache dar. Zusammen mit den Auswirkungen selektiver Kräfte bei der Stichprobenbildung, vor allem durch die hohen Anforderungen an die Buchhaltungsform, lässt sich insgesamt folgern, dass die heute berechneten und publizierten landwirtschaftlichen Einkommen der Referenzbetriebe gegenüber den Werten der Grundgesamtheit um 3% bis 8% zu hoch sind. Bei anderen Kennzahlen fallen die Verzerrungen grösser oder auch geringer aus.

Eine aus statistischer Sicht erforderliche Zufallsauswahl würde unter den aktuellen Rahmenbedingungen zu so tiefen Antwortraten führen, dass im Wesentlichen mit ähnlichen Verzerrungen zu rechnen wäre, wie sie zur Zeit an der Stichprobe der Referenzbetriebe beobachtet werden. Eine grobe Schätzung zeigt demgegenüber auf, dass mit den Massnahmen gemäss „Methode 2000“ über 80% der Einkommensdifferenz zwischen den früheren Testbetrieben und dem vermuteten Mittelwert der neuen Grundgesamtheit korrigiert wird. Somit können diese Massnahmen als wirksam und aufgrund des geringen Umsetzungsaufwandes auch als effizient beurteilt werden.

Es besteht jedoch weiteres Verbesserungspotenzial. Die kurzfristig umsetzbaren Vorschläge zur Verbesserung der Schätzqualität betreffen eine Dynamisierung der Schwellen zur Abgrenzung der Grundgesamtheit, Anpassungen beim Gewichtungssystem und bei der Auswahl untervertreter Betriebsgruppen, die Prüfung eines neuen Masses für die Betriebsgrösse, eine Angleichung von angestrebter Grundgesamtheit und Auswahlgesamtheit sowie eine vereinfachte regelmässige Verzerrungsschätzung anhand ausgewählter Hilfsvariablen.

Etwas aufwändiger in der Umsetzung sind Vorschläge zur Vertiefung der vorliegenden Verzerrungsschätzung, Erweiterungen des Gewichtungssystems mit zusätzlicher Kalibrierung von Hilfsvariablen sowie die Einführung der Zufallsauswahl für Teilstichproben.

Die vorliegende Arbeit stellt selektionsbedingte Verzerrungen in den Vordergrund. Eine Optik für den Gesamtfehler ist jedoch erforderlich, wenn konzeptionelle Alternativen zum aktuellen Erhebungssystem zu beurteilen sind. Aufbauend auf den hier analysierten Selektionsfehlern muss dabei den Wechselwirkungen mit Zufallsfehlern und Messfehlern grosse Beachtung geschenkt werden.

## Summary

Until 1998, the Farm Accountancy Data Network in Switzerland was based on a sample of so-called test farms. This sample was deliberately selected to include farms characterised by above-average economic results. However, one of the objectives of the redrafted legal mandate of 1999 was to illustrate the economic situation in agriculture with a sample of representative reference farms. Starting from existing data, the aim was to achieve this objective with the “Method 2000” package of measures. The most important methodical changes comprise the delimitation of the population, the development of a new farm typology, the introduction of a stratified selection plan and weighting of the individual farm results.

However, statistical conclusions from the sample to the population are problematic because there is no random selection of reference farms. Considerable bias must be expected for the averages estimated today. The quality of these estimates is assessed by several methodically different approaches. The results show that due both to the inadequate representation of specialist crops and pigs/poultry farms and to the exclusion of farm associations from the reference farms, today’s estimated family farm incomes show a negative bias in the order of 2% relative to the target population, i.e. the current estimate is too low. A calibration model for the adjustment of biased auxiliary variables (area, age, education, organic farms) does, however, indicate a positive bias of agricultural income of between 5% and 6%. The above-average education of the farm managers in the current sample constitutes a primary cause. Together with the impact of selective forces in sampling, due mainly to the high standards demanded of the accounting system, it can on the whole be concluded that the family farm incomes of the reference farms calculated and published today are between 3% and 8% too high relative to the figures of the population. The bias is greater or also smaller for other key figures.

In the current general framework, a random selection necessary from the statistical point of view would lead to such low response rates that we could basically expect the bias observed today. However, a rough estimate shows that, with the measures under “Method 2000”, over 80% of the income difference between the previous test farms and the assumed average of the new population is adjusted. These measures can therefore be judged effective as well as efficient owing to the low implementation costs.

However, there is further potential for improvement. The proposals for improving estimation quality, suitable for short-term implementation, relate to a dynamisation of thresholds for the delimitation of the population, adjustments in the weighting system and in the selection of under-represented farm groups, the examination of a new measurement of farm size, an adjustment of target population and frame population, and a simplified regular estimation of bias based on selected auxiliary variables.

Associated in part with higher implementation costs are proposals for bias assessment in greater depth than is currently the case, extensions of the weighting system with additional calibration of auxiliary variables, and the introduction of random selection for sub-samples.

The present paper gives priority to selection-related bias. However, total survey error needs to be taken into account if conceptual alternatives to the current survey system are to be assessed. Building on the selection errors analysed here, careful attention must at the same time be paid to interactions with sampling errors and measurement errors.

## Résumé

Jusqu'en 1998, le Dépouillement centralisé des données comptables en Suisse a été effectué avec un échantillon d'exploitations-témoins sciemment sélectionnées sur la base de leur résultat économique positif. Le mandat légal reformulé en 1999 a pour but de représenter la situation économique de l'agriculture suisse au moyen d'un échantillon d'exploitations de référence représentatives. Ce nouvel objectif est poursuivi sur la base des données existantes et au moyen du programme de mesures „Méthode 2000“. Les principaux changements méthodologiques concernent la délimitation de la population-cible, le développement d'une nouvelle typologie des exploitations, l'introduction d'un plan de sélection stratifié et de la pondération des résultats des exploitations individuelles.

En raison de l'absence d'une sélection aléatoire des exploitations de référence, il est cependant problématique de tirer des conclusions statistiques pour la population-cible sur la base de l'échantillon. Il faut s'attendre à ce que les moyennes estimées comprennent des distorsions considérables. La qualité des estimations a été vérifiée avec un nombre de méthodes différentes. Les résultats montrent que, due à la sous-représentation des exploitations de cultures spéciales et des exploitations de transformation et à l'exclusion des communautés d'exploitations des exploitations de référence, les revenus agricoles estimés à l'heure actuelle présentent une distorsion négative d'environ 2 % par rapport à ceux de la population-cible, c'est-à-dire que les valeurs estimées sont trop basses. Le modèle de calibration utilisé pour corriger les variables auxiliaires distordues (surface, âge, niveau de formation, exploitations biologiques) indique en revanche une distorsion positive du revenu de 5 % à 6 %, ce qui est principalement dû, dans l'échantillonnage actuel, au niveau de formation du chef d'exploitation supérieur à la moyenne. Tenant compte des effets résultant des forces sélectives lors de la formation de l'échantillon (notamment dus aux exigences à respecter par le système de comptabilité), on peut globalement en conclure que les revenus agricoles calculés et publiés à l'heure actuelle sont de 3 à 8 % trop élevés par rapport à ceux de la population-cible. Pour d'autres paramètres calculés, les distorsions sont plus élevées ou plus basses.

Une sélection aléatoire telle qu'elle serait nécessaire d'un point de vue statistique conduirait à des taux de réponse si basses dans les conditions-cadres actuelles qu'il faudrait s'attendre à des distorsions dans l'ordre de grandeur actuelle. Selon une estimation grossière, les mesures de la „Méthode 2000“ permettront de corriger plus de 80 % de la différence entre le revenu des anciennes exploitations-témoins et la moyenne estimée pour la population-cible. Les mesures peuvent donc être considérées comme efficaces et économiques étant donné leur dépenses peu élevées.

D'autres améliorations sont cependant possibles. Les mesures suivantes peuvent être mises en œuvre à court terme pour améliorer la qualité d'estimation: dynamisation des seuils pour la

délimitation de la population-cible, adaptation du système de pondération et de la sélection des groupes d'exploitations sous-représentés, évaluation d'une nouvelle unité pour mesurer la taille des exploitations, adaptation de la population-cible et du cadre de sondage ainsi qu'une estimation simplifiée régulière des distorsions sur la base de variables auxiliaires sélectionnées.

D'autres mesures d'amélioration impliqueraient de plus grandes modifications: approfondissement de l'estimation actuelle des distorsions, élargissement du système de pondération par un calibrage supplémentaire des variables auxiliaires ainsi que l'introduction d'une sélection aléatoire pour l'échantillonnage partiel.

Le présent travail porte essentiellement sur les distorsions occasionnées par la sélection. Pour développer et évaluer des alternatives ou compléments au système de relevé actuel, il est indispensable de prendre en compte l'erreur totale afin de pouvoir satisfaire les exigences complexes du Dépouillement centralisé des données comptables. Pour ce faire, il faut se baser sur les erreurs de sélection analysées ici, tout en tenant compte des interactions avec les erreurs aléatoires et les erreurs de mesure.