Economic Modelling 38 (2014) 238–245

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.elsevier.com/locate/ecmod

# Model structure or data aggregation level: Which leads to greater bias of results?

Martina Brockmeier *, Beyhan Bektasoglu

*Institute for Agricultural Economics and Social Sciences in the Tropics and Subtropics, Chair of International Agricultural Trade and Food Security (490b), University of Hohenheim, Institut 490b, 70593 Stuttgart, Germany*

ARTICLE INFO

*Article history:*
Accepted 8 January 2014
Available online 31 January 2014

*JEL classification:*
F13
D58
Q17

*Keywords:*
System of linked models
Model structure
Aggregation level
False competition
Tariff aggregation
GTAP

ABSTRACT

The aim of this paper is to provide a first step toward a systematic sensitivity analysis of a system of linked models. We focus on two fundamental characteristics: the model structure and the data aggregation level. Employing the Global Trade Analysis Project (GTAP) framework, we combine the general equilibrium (GE) and partial equilibrium (PE) versions of the GTAP model, each of which is then run with a highly aggregated and a highly disaggregated version of the GTAP database. Based on this experimental setting, we quantify the biases resulting from the data aggregation, the model structure and the interaction of these two model characteristics. We conclude that data aggregation as well as the related false competition and tariff averaging influence the results significantly more than the model structure, whereas the bias stemming from the interaction of the two model characteristics is negligible.



## 1. Introduction

Agricultural policy has become increasingly complex over the last century. Reforms and new areas of concern have led to the implementation of complex measures that influence the food and agricultural sector, from global markets to the level of individual farms. As a result, quantitative studies of food and agricultural policy have become a greater challenge that calls for increasingly comprehensive analytical tools and a modeling framework that represents the food and agricultural sector at the global, national and farm levels.

It is obvious to modelers that there is no “one-size-fits-all model” to analyze such widespread food and agricultural research questions. An integrated model that is fully consistent at all levels of data aggregation is not yet available because of computational capacity constraints. The currently preferred approach is to utilize the comparative advantages of different types of models and combine them in a strategically useful way to more accurately represent the micro and macro aspects of the food and agricultural sector. Consequently, in recent years, we have observed an increase in the development and application of systems of linked models.

The usefulness of linked models in research and particularly in policy advice is discussed at length in the literature. This discussion reveals that systems of linked models can be characterized according to (1) differences in data, parameters, policy instruments or other model structure; (2) the existence of formal or informal linkages between linked models; (3) the direction in which results are transferred between models (top down, bottom up or iteration); and (4) the approach that is employed to aggregate or disaggregate results that are transferred to the next model in the system. Systems of linked models are built with different combinations of the aforementioned options, although the choice of a specific combination significantly influences the outcome of the analysis.

To the best of our knowledge, no existing study provides quantitative estimates of the bias in results related to the approach that is used to build a system of linked models. Systematic sensitivity analysis has certainly not been used for this purpose. The predominant approach in the literature involves comparing the results of a partial equilibrium (PE) model constructed with disaggregated data to the results of a general equilibrium (GE) model developed with aggregated data. In other words, two fundamental characteristics are changed simultaneously in this comparison. This procedure clearly does not allow for the effects of different model structure to be distinguished from the effects of data aggregation. Moreover, the existing papers do not derive conditions for the optimal interaction of the employed models, including one or several characteristics of the approaches to link models mentioned above. For example, would it be most effective to align the sector disaggregation of two adjacent models or to develop a more compatible model structure to obtain unbiased results within the system of linked models?

* Corresponding author at: University of Hohenheim, Institut 490b, 70593 Stuttgart, Germany. Tel.: +49 711 459 23392; fax: +49 711 459 23762.
*E-mail addresses:* martina.brockmeier@uni-hohenheim.de (M. Brockmeier), beyhan.bektasoglu@uni-hohenheim.de (B. Bektasoglu).

0264-9993/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.econmod.2014.01.003

We contribute to the existing literature in several ways. This paper is a first attempt to conduct a systematic sensitivity analysis of systems of linked models. To illustrate the procedure, we use a simplified experimental setting by focusing on the differences between sector disaggregation and the structures of two adjacent models. In particular, we first concentrate on aggregation problems that arise as a result of the transfer of results between linked models that are aggregated differently. Second, we account for differences in the structures of PE and GE models. Our purpose is to extend beyond the well-known argument that PE models are not suited to examine policy shocks outside of their predefined domain (e.g., the use of agricultural PE models to analyze shocks in non-agricultural sectors), whereas GE models generally do not capture the necessary details of the sector of immediate interest. Rather, we seek to uniquely quantify the bias that is introduced in a system of linked models by transferring results from one model to another within one sector; an example would be the effects of global agricultural trade liberalization on global (GE model), national and farm levels (PE model). Additionally, we wish to account for the interaction of the aforementioned characteristics by examining all possible combinations of the four model features (i.e., disaggregated and aggregated data and PE and GE model structures). For this purpose, we require a tool that is sufficiently flexible to address these four characteristics, that clearly distinguishes them from one another and that therefore enables us to quantify the deviation of the outcomes of experiments from a predefined reference situation (e.g., a disaggregated GE model). These requirements are satisfied by the Global Trade Analysis Project (GTAP) framework, which we apply in this paper.

This paper is organized as follows. Section 2 briefly reviews the literature to systemize the results of papers comparing systems of linked models with respect to selected variables, namely, trade flows, prices and output. Section 3 introduces the GTAP framework and its adaption to the empirical requirements of this paper. In Section 4, we conduct a systematic sensitivity analysis to determine whether model structure or data aggregation is more important to ensure unbiased results in systems of linked models. A conclusion summarizes the main results.

## 2. Literature review

PE and GE models have long been used to analyze trade liberalization. Although many authors employ either PE or GE standalone models, an increasing number of researchers apply PE and GE models within a linked system. To the best of our knowledge, few papers have attempted to compare the results of linked models and used this exercise to deduce implications for how to address systems of linked models (e.g., Gohin and Moschini, 2006). The following literature review summarizes the findings of these papers by focusing on a selective number of variables representing trade, welfare, price and output effects (see Table 1). Additionally, we distinguish effects resulting from aggregation from those resulting from model structure (PE and GE models).

In the upper part of Table 1, we document the findings of papers in which the authors compare models with different levels of aggregation. For instance, Charteris and Winchester (2010) compare a GE model with a disaggregated dairy sector and joint production to an aggregated GE model without joint production. The authors argue that trade liberalization has different effects that result from the aggregation level of the database employed by the models and the protection structure. Greater variation in tariff rates and higher elasticities of substitution between disaggregated commodities increase these differences. Grant et al. (2007) combine a highly disaggregated PE and a GE model to analyze tariff rate quotas in the US dairy sector. The authors compare the performance of the linked PE/GE model with that of the standalone GE model. Their results demonstrate that the terms of trade effects are poorly predicted by the standalone GE model, whereas the welfare effects are within a similar range. Narayanan et al. (2010a) analyze the effects of different model structures and data aggregation levels using the complete liberalization of the Indian automotive industry as an example. The authors compare the outcome of three different models, namely a disaggregated PE model, an aggregated GE and a model that links these PE and GE models.[1] The results of their study reveal that a higher level of aggregation leads to larger trade effects. Nielsen (1999) also investigates the effect of data aggregation and model structure on EU enlargement with the use of six models with different closures (five PE models and one GE model). In contrast, her results demonstrate that the trade effects are generally smaller when an aggregated database is used in the simulation.

These studies also exhibit similar results in terms of welfare, price and quantity effects (see Table 1). In most cases, aggregation causes smaller changes in output and price levels as well as smaller welfare gains (Nielsen, 1999; Grant et al., 2007; Narayanan et al., 2010a; Charteris and Winchester, 2010). But Grant et al. (2007) find only minor differences between models with respect to welfare results.

The lower part of Table 1 presents a comparison of the model results with respect to model structure. Wailes and Morat (2005) employ both general and spatial partial equilibrium models to quantify the effects of the Central America Free Trade Agreement (CAFTA) on the US rice industry. The authors conclude that the trade effects that are quantified using the GE and the PE models generally point in the same direction and exhibit only small differences resulting from the level of data aggregation and model structure. Furthermore, Narayanan et al. (2010a) demonstrate that GE model estimates of changes in aggregate imports are generally larger than those obtained in PE models. Additionally, Nielsen (1999) demonstrates that the GE model structure leads to smaller changes in aggregate exports and imports.

With respect to price and output effects, some authors argue that model structure causes only minor differences (Hertel, 1992; Wailes and Morat, 2005; Gohin and Moschini, 2006). Narayanan et al. (2010a) find that quantity changes are larger but price changes are smaller in GE models. Conversely, Nielsen (1999) demonstrates that both price and quantity changes are generally smaller in GE models. Furthermore, a study by Gylfason (1995) comparing the costs of agricultural supports calculated by various models under different equilibrium conditions reveals that the GE model yields higher cost estimates than the PE model.

In the last column of the lower part of Table 1, we present the effects of the model structure on welfare results. In previous studies, Tokarick (2003) and Gohin and Moschini (2006) report a larger welfare effect when using a GE structure. In a study measuring the effects of distortions in agriculture trade using different model structures, Tokarick (2003) reports that the GE structure yields larger welfare effects through income, demand and price mechanisms. Additionally, the author shows that the welfare effects of liberalization depend on the model structure and primarily result from higher efficiency and not from the terms of trade effect. Furthermore, in a meta-analysis that is designed to compare model structures, Hess and Cramon-Taubadel (2007) report findings that contrast with those in previous papers; namely, PE models and aggregation yield greater welfare changes.

## 3. Modeling framework and methodology

The analysis in this paper utilizes the GTAP modeling framework. The standard GTAP model follows the typical structure of a static multi-regional general equilibrium model. As such, this model exhibits an economy-wide representation of each region or country, including the linkages between the farming, agribusiness, industrial and service sectors. Bilateral trade flows are represented by the Armington assumption, and a non-homothetic constant difference of elasticity (CDE)

[1] The choice of model structure and data aggregation in Narayanan et al. (2010a) simultaneously reflects differences resulting from model structure and data aggregation. Hence, the authors are unable to isolate the effects of different model structures and levels of data aggregation. We use their results, which are reported in Table 3 (p. 763), to obtain findings that are suitable for our comparison in Table 1.