Chapter 20

Aggregation and calibration

Betina Dimaranan, Thomas Hertel, Robert McDougall

In the previous chapter we discussed how the final version 3 GTAP data base was assembled. This data base is extremely large. For example, just one of the arrays associated with intermediate usage of commodities is of dimension: 37 x 37 x 30 = 41,070 elements. For this reason, the resulting model is also extremely large. Indeed for most researchers, a model based on the fully disaggegated data base is simply too demanding of computer memory and disk space. As a result, most GTAP users work with aggregations of the data base. Procedures for aggregation are discussed in section 20.1.

Once the data base and elasticities have been aggregated, there remains the problem of calibration. Fortunately, the extent of this problem is minimized by solving the non-linear GTAP model via its linearized representation. This approach has the advantage of not requiring calibration of the so-called “shift parameters” in the production and utility functions, as they are embedded in the cost and expenditure shares which appear in the differential form of the model. In the case of the CES functional form, where the shares and the constant substitution elasticity parameter are all that appear, no further calibration is required. However, in the case of the CDE functional form used to represent private household demands, a further calibration step is required. This is discussed in section 20.2.

20.1 Aggregation Procedures

Aggregation of value flows in the benchmark data is involves the simple summation of appropriate component elements. For example, disaggregate consumer expenditures for agricultural and processed food products are summed to obtain total consumer expenditures for the aggregate category, food. Similarly, when aggregating regions, consumer expenditures for each commodity are summed across disaggregate regions. For variables that are source-specific (e.g., trade flows), the regional aggregation also aggregates across these sources.

Aggregation of behavioral parameters of the model is more involved than aggregation of value flows. The relevant behavioral parameters are the uncompensated own-price and income elasticities of demand, the elasticities of substitution in value-added, and the elasticities of substitution among foreign goods and between domestic and foreign goods. The own-price and income elasticities of demand vary by commodity and region, whereas the substitution elasticities vary only by commodity. Rather than simply computing arithmetic means of the disaggregate parameters, we compute their expenditure share-weighted means.

In the case of consumer demand elasticities, we proceed as follows:

— We ensure that the Engel aggregation condition holds for the disaggregate data and parameters by proportionally adjusting all income elasticities as required.

— We apply the Slutsky condition to derive *compensated* own-price elasticities based on the GTAP data for *uncompensated* own-price elasticities and the income elasticities from the previous step.

— We obtain expenditure share-weighted means of the disaggregate income and *compensated* own-price elasticities for the aggregation in question.

— We again make sure that the Engel aggregation condition holds for the aggregated data and income elasticities by proportionally adjusting all income elasticities.

Since the elasticities of substitution are not region-specific, only aggregation across commodities is necessary. Hence, the share weights used in commodity aggregation are global values. For the elasticities of substitution in value-added, sectoral payments to land, labor, and capital are summed to obtain total value-added by sector, which is the weight used in the aggregation. For the elasticities of substitution among foreign goods and between domestic and foreign goods, total expenditures for imports and total expenditures for imported and domestic commodities are calculated. As above, share weights are then calculated for the disaggregate members of each aggregate commodity and used to weight the disaggregate substitution elasticities prior to summation.

Table 20.1 reports the parameter file used in the 3x3 aggregation taken from Hertel *et al*. It begins with the calibrated CDE parameters, SUBPAR and INCPAR, which are commodity- and region-specific (see next section for more details). These are followed by the region-generic elasticities of

Table 20.1 Parameters for the 3x3 Aggregation

```
! The matrix SUBPAR(%1,%2) with %1 in IND, %2 in REG.
 !%2= USA         EU               ROW
   0.866678       0.897081         0.967472       !%1=food
   0.239291       0.317116         0.669294       !%1=mnfcs
   9.999990E-03   9.999990E-03     9.999990E-03   !%1=svces

 ! The matrix INCPAR(%1,%2) with %1 in IND, %2 in REG.
 !%2= USA         EU               ROW
   0.133962       0.160812         0.331707       !%1=food
   0.793119       0.789978         0.838876       !%1=mnfcs
   1.14379        1.26177          1.29555        !%1=svces

 ! Values of ESUBD(IND) - an array of size 3
 !--------------------
 !   food          mnfcs             svces
    2.39901     2.79556        1.94365

 ! Values of ESUBM(IND) - an array of size 3
 !--------------------
 !
 !   food          mnfcs             svces
   4.63905      6.08810        3.91673
 !
 ! Values of ESUBVA(IND) - an array of size 4
 !---------------------
 !
 !   food          mnfcs             svces
    0.789314    1.21992              1.38946
 !   CGDS
    0.000000
 !
 ! Values of ETRAE  (a single number)
 !---------------
 !
   -1.00000
 !
 ! Values of RORFLEX(REG) - an array of size 3
 !----------------------
 !
 !   USA         EU             ROW
    10.0000     10.0000        10.0000
 !
 ! Values of RORDELTA  (a single number)
 !------------------
 !
    0.000000
```

Source: *Global Trade Analysis: Modeling and Applications*, Thomas W. Hertel (ed) Chapter 5, p. 156.

substitution between domestic goods and imports (ESUBD), between imports from different sources (ESUBM), and among the components of value-added (ESUBVA). The next entry, ETRAE, specifies the elasticity of transformation between alternative uses of the sluggish endowment commodities – in this case, farmland. The final sets of parameters are relevant to the interregional allocation of investment. RORFLEX, the flexibility of investment allocation, is permitted to vary by region. It is operational only when RORDELTA is equal to 1. In this parameter file *RORDELTA* = 0, so that the global banking sector's regional investment portfolio is fixed.

20.2 The CDE Functional Form

For GTAP, we choose the CDE functional form for the private household's expenditure function. The CDE expenditure function was introduced by Hanoch (1975), who discussed models that were more general than the CES but less general than a flexible functional form, for example, the translog.[3] The CDE is based on the assumption of implicit additivity. which, in the case of N commodities, constrains the symmetric $N \times N$ matrix of elasticities of substitution to depend on only N parameters. The CDE also allows for a richer representation of income effects in the demand system. In particular, marginal budget shares may vary with expenditure levels. Section 20.2.1 gives a brief overview of the CDE form. Section 20.2.2 discussed the calibration problem in detail.

20.2.1 Theoretical Development of the CDE

In general, an expenditure function can be represented in the following manner:

$$E = G(p,u) = \{\min p^{/} x : f(x) \geq u\}, \tag{20.1}$$

where p and x are N-dimensional vectors of prices and demands, u is utility, and E is minimum expenditure. The function $f(\cdot)$ represents utility, and $G(\cdot)$ is the minimum expenditure function. Function $G(\cdot)$ is homogeneous of degree 1 in prices, allowing the following normalization of prices and expenditure by minimum expenditure:

$$G(E^{-1}p,u) = G(z,u) \equiv 1, \tag{20.2}$$

where the z's are the normalized prices. To obtain the CDE expenditure function, Hanoch (1975) restricts the number of substitution effects to N by imposing additivity in the normalized prices. The implicit function proposed takes the form:

$$G(z,u) \;=\; \sum_{i=1}^{N} B_i u^{e_i b_i} z_i^{\,b_i} \;\equiv\; 1\,, \tag{20.3}$$

where the b_i's are the N parameters, which determine substitution possibilities among commodities in consumption; the e_i's are N expansion parameters, which appear owing to nonhomotheticity in consumption; and the B_i's are scale parameters necessary to specify the function. It is required that $B_i > 0$ and $e_i > 0$, and $b_i < 1$, with either $0 < b_i < 1$ or $b_i < 0$ for all i (Hertel et al. 1991).

If the substitution parameters are rewritten as $\alpha_i = 1 - b_i$, the Allen partial elasticities of substitution can be expressed as:

$$\sigma_{ij} \;=\; \alpha_i + \alpha_j \;-\; \sum_k s_k \alpha_k \;-\; \frac{\delta_{ij}\alpha_i}{s_i}\,, \tag{20.4}$$

where $\delta_{ii} = 1$, and $\delta_{ij} = 0$ for $i \neq j$, and the s_i's are expenditure shares. The name *constant difference in elasticities* arises due to the fact that the difference between the elasticities of substitution σ_{ij} and σ_{ih} is invariant to index i (Hertel et al. 1991)

$$(\sigma_{ij} - \sigma_{ih}) = (\alpha_j - \alpha_h) \tag{20.5}$$

The expressions for income elasticities of demand are:

$$\eta_i = \left(\sum_k s_k e_k \right)^{-1} \left[e_i (1-\alpha_i) + \sum_k s_k e_k \alpha_k \right] + \left(\alpha_i - \sum_k s_k \alpha_k \right)\,, \tag{20.6}$$

and the expressions for compensated own-price elasticities of demand are:

$$-v_{ij} \;=\; -s_i \left[2\alpha_i \;-\; \sum_k \alpha_k s_k \;-\; \alpha_i / s_i \right]. \tag{20.7}$$

20.2.2. Calibration of the CDE

The income and uncompensated own-price elasticities of private household demand used in GTAP derive from a variety of sources, *none* of which involve econometric estimation using the CDE form. This is a common difficulty when choosing elasticities for AGE models. That is, the functional forms used in the econometric studies from which estimates are drawn do not correspond to the functional forms specified in the model. Furthermore, the sample period used in estimation may not include the benchmark equilibrium year. Finally, econometric studies are based on consumer goods prices, whereas AGE models typically evaluate consumption at producer prices. These divergences cause difficulty, as not all sets of income and own-price elasticities will be representable via the CDE functional form, evaluated at 1992 producer prices.

The problem can be seen clearly for the consumption budget shares and own-price elasticities shown at the top of Table 20.2. These are derived from a three-region, three-commodity aggregation of GTAP (Huff et al). The α_i parameters derived by solving equation (20.7) for each region are also given in Table 20.2. Plainly, none of these sets of values satisfies the requirement that $0 < \alpha_i < 1$ for all *i*. Therefore, a procedure must be formulated that, for any set of *target* income and own-price elasticities, determines values of the CDE parameters that yield *model* income and own-price elasticities sufficiently close to the target income and own-price elasticities. This section describes such a calibration procedure for determining CDE parameters.

It is evident from the outset that certain aspects of this calibration procedure, such as the choice of a measure of "closeness," will be, to some extent, arbitrary. Another element of choice is the weight to be given to deviations of own-price elasticities from target values as compared with income elasticities. In the case of the CDE functional form, a "natural" choice in this regard is suggested by equations (20.6) and (20.7), relating elasticities to CDE parameters. The own-price elasticities are determined by the α_i parameters, while the income elasticities depend on both α_i and e_i parameters. Thus the procedure adopted was to determine α_i's to ensure closeness of the own price elasticities implied by equation (20.7) with target elasticities, and to use these α_i's in equation (20.6) when determining an appropriate set of e_i parameters.

The calibration procedure for CDE parameters entails solving two constrained minimization problems. The first determines the CDE α_i parameters by seeking a good fit to the target own-price elasticities. The second, taking the α_i values from the first as given, determines the e_i parameters by seeking a good fit to the target income elasticities. The constraints imposed are of two types. First are those on CDE parameters required for the regularity of the CDE form. Second is the requirement that income elasticities are greater than (less than) or equal to 1 if the corresponding targets are greater than (less than) 1. This latter constraint ensures that goods that were superior (inferior) as represented by the target elasticities remain so after calibration.

The constrained minimization problem for determining the CDE α_i parameters can be represented formally as: given ν_{ii}^{target}, choose ν_{ii}, and α_i to minimize:

$$\sum_i \nu_{ii}[\log_e(\nu_{ii}/\nu_{ii}^{target}) - 1] \tag{20.8}$$

subject to $0.01 \leq \alpha_i \leq 0.99$, and equation (20.7), where ν_{ii}^{target} = target-compensated own-price elasticities; ν_{ii} = compensated own-price elasticities arising from calibration; α_i = CDE substitution parameters arising from calibration.

The regularity requirement that $0 < \alpha_i < 1$ has been replaced by the inequality constraint shown. Otherwise the problem may not have a solution if the objective function attained its minimum on the boundary of the region $\{\alpha_i: 0 < \alpha_i < 1 \text{ for all } i\}$. The objective function is strictly convex, and the feasible region convex, so a unique minimum exists.

The constrained minimization for determining the CDE e_i parameters can be represented formally as: given η_i^{target}, choose η_i, and e_i to minimize:

$$\sum_i (\eta_i - \eta_i^{target})^2 \tag{20.9}$$

subject to equation (20.6), $0.01 \leq e_i$, $\sum_i s_i \eta_i = 1$, and $(\eta_i -1)*(\eta_i^{\text{target}} -1) \geq 0$,

where η_i^{target} = target income elasticities; η_i = income elasticities resulting from calibration; and e_i = CDE expansion parameters resulting from calibration.

The regularity requirement that $e_i > 0$ has been replaced by the inequality constraint shown, for the same reason that the regularity condition on the α_i was altered in the first minimization. It should be noted that the second constraint ensures that the income elasticities derived as the solution to this minimization will satisfy the Engel condition. The third constraint ensures that the calibrated income elasticities lie on the same side of one as the target elasticities.

Although many software packages exist in which these constrained minimizations could be formulated and solved with ease, they have been implemented in GEMPACK. This has been done to ensure that GTAP users do not require access to any other software in order to perform the entire range of operations, from data aggregation to printing of simulation results.

Finally, a important warning. It is key for users to note that the calibration process inevitably results in elasticities which depart from their target values. The degree to which this occurs will vary by aggregation. In many cases the target and fitted income elasticities of demand are quite close. However, users have also reported some instances where they depart significantly. If the private household's income and price responsiveness is important to the problem at hand, the user would be well-advised to spend some time scrutinizing the fitted values before proceeding.

Table 20.2 A 3X3 Example of the CDE

	USA	EU	ROW
		Consumption Shares	
Foods	0.082048	0.151911	0.212457
Manufactures	0.173766	0.200349	0.198768
Services	0.744186	0.64774	0.588775
		Own-Price Elasticities	
Foods	-0.172441	-0.177254	-0.149364
Manufactures	-0.547482	-0.501926	-0.31344
Services	-0.428958	-0.418842	-0.282822
		α_i Parameters Derived by Solving Equation (20.7)	
Foods	-0.40	-0.23	-0.10
Manufactures	-0.80	0.09	0.20
Services	8.51	3.46	1.60

References

Hanoch, G. (1975) "Production and Demand Models with Direct or Indirect Implicit Additivity," *Econometrica* 43:395–419.

Hertel, T. W., E. B. Peterson, P. V. Preckel, Y. Surry, and M. E. Tsigas (1991) "Implicit Additivity as a Strategy for Restricting the Parameter Space in CGE Models," *Economic and Financial Computing* 1(1):265–289.

Huff, K., K. Hanslow, T.W. Hertel, and M.E. Tsigas (1997) "GTAP Behavioral Parameters." In Thomas W. Hertel (ed), *Global Trade Analysis: Modeling and Applications*, Chp. 4, pages 124-148.