(nb: incomplete: more to add later)

Aggregating CES elasticities for CGE models

Mark Horridge, Centre of Policy Studies,
Victoria University, Melbourne, Australia

THURSDAY, JUNE 14 3:00-4:30pm (15:00-16:30) **Organized Session #35 302A** (Level 3)

Abstract

It is expensive to prepare a database for a large CGE model -- so we should plan that the database will be used for many (perhaps unanticipated) purposes. For this and other reasons it is wise to construct a database with as much regional and sectoral disaggregation as possible. But a model with too many sectors and regions is annoyingly slow to solve. Thus it is common practice to aggregate a large 'master' database before use, in a way that reflects the policy focus of a particular set of simulations. However, we may wonder if aggregation greatly affects results. Are there choices we can make, when aggregating, which would cause results computed with aggregated data to more closely resemble those with disaggregate data?

A CGE database consists mainly of matrices of flow values, and elasticities (mostly pertaining to CES nests). Having defined many-to-one mappings from the (many) old sectors/regions to the (fewer) new sectors/regions, it is easy to aggregate the flows matrices by simply adding.

Aggregated elasticities are usually constructed as weighted averages of the disaggregate elasticities -- using as weights the flow values associated with each elasticity (that is, the total cost of inputs to that nest). However, we present examples where this simple weighted averaging yields odd results.

We focus on the case where a number of users (or nests) each combine (using the CES) the same set of inputs (but with different cost shares). Seeking a better method of averaging elasticities, we propose as a criterion that aggregated CES elasticities imply a local own-price substitution response that is close to the average of the own-price responses in the corresponding disaggregated nests. We see that elasticity aggregation bias arises from aggregating nests, rather than aggregating inputs to nests.

We develop some simple formulae, which imply that the aggregate elasticity should be K (0<K<1) times the disaggregate elasticity where K is smaller as the variance of disaggregate cost shares is larger. We show how such formulae could be applied to the GTAP model.

The formulae suggest that aggregation of sectors/regions should in general reduce CES elasticities, by an amount depending on the variance of disaggregated input shares within a specific aggregated CES nest. However, most CGE models specify that a number of CES nests share the same CES elasticity. For example, in the GTAP model each of 140 regions contains 60 users (57+C,I,G) of Textiles, each with unique import/domestic shares modelled by a CES nest. Yet the standard model requires that all 8400 CES nests share a single worldwide substitution elasticity (ESUBD=3.75). We propose that allowing regional subscripts on elasticities would allow for better elasticity aggregation, even if, in the disaggregated master database, elasticity estimates were uniform across regions.

A related problem arises when a CGE model use external elasticity estimates which were estimated using a very different level of aggregation to the CGE model. For example, a third party may have estimated a CES capital-labour substitution elasticity for "manufacturing" . Should we apply that value to all 12 manufacturing sectors in our model? The "K approach" suggests that disaggregate elasticities should be larger. But elasticity estimates vary widely for several reasons, so if K>0.95, we should probably not worry much about aggregation bias. To gain a sense of the scale of the problem, we plan to measure K (aggregation bias) using values from the GTAP and TASTE datasets.

Aggregating CES elasticities for CGE models

Mark Horridge, Centre of Policy Studies, Victoria University, Melbourne[1]

1 Introduction[2]

To prepare a new database for a large CGE model is an expensive exercise -- usually too expensive to justify for one or a few simulations. To make it worthwhile, we should plan that the database will be used many times, probably for a variety of quite different experiments -- which cannot all be anticipated now. Hence, we should attempt to preserve as much sectoral and regional detail as possible, in case this is useful later. In this way the expense of building the database may be spread over a number of projects. These points are reinforced for a database, like GTAP's, which is created and used by a group of partners with diverse areas of interest.

Yet the downside of a large database (GTAP distinguishes 140 regions and 57 sectors) is that a model based directly on that database may be annoyingly slow to run, especially if the model is to be solved for several time periods. So it is usual to aggregate the database before use, by combining sectors and regions. The system of grouping sectors and regions will be devised afresh for each project, in such a way that detail relevant to the current task is preserved, and less relevant detail aggregated away. Thus, to analyse an East Asian trade deal we would retain China, Japan, Korea, etc as separate regions, while perhaps grouping non-Asian counties into a few continents. A 25 region/40 sector model might solve 100 times quicker than one using the full-size data.

Apart from speeding up simulations, modelers may be tempted to aggregate for two less worthy reasons:

- Flows which are tiny, zero or negative can cause problems in CGE models; aggregation may mask a lot of these problems. However, there is usually a way to fix such problems properly, rather than hiding them.
- Reducing the dimensions of the model may make it easier to fit tables of results onto a page, or even onto a Powerpoint slide. However, it should be quite easy to add model code that computes more concise results tables for groups of sectors or regions.

The question then arises, does aggregation alter or bias results in any systematic way? Are there procedures to reduce such bias?

An aggregation scheme applies to a set of categories such as sectors, regions, household types or primary factors. It usually consists of a series of many-to-one mappings so that, for example, each original sector is wholly contained within one new sector (which consists of one or more old sectors).

The original CGE database will consist of flow matrices, with region/sector/etc dimensions, and elasticities (often CES elasticities). For the aggregated database, flows may be simply computed by adding up blocks of the original data matrices. Usually, each original elasticity is assigned a weight; and each aggregated elasticity is a weighted average of several original elasticities[3].

However, this standard method of aggregating elasticities can yield paradoxical results, as the next example shows.

---

[1] Mark Horridge, Centre of Policy Studies, Victoria University, Melbourne, mark.horridge@gmail.com

[2] This paper benefits from discussions with Tom Rutherford in Melbourne during the 2015 GTAP conference.

[3] There might be another step to 'normalize' the aggregated elasticities; for example, LES marginal budget shares might be scaled to add to one.

1.1 An extreme example of aggregation bias

Consider two sectors, each using a CES production function to combine capital and labour -- but in quite different proportions (see Table 1). For both industries σ=1. What would happen to total employment if wages rose by 1%, holding sector outputs constant?

Further below (in section 2.1), we explain that the own-price elasticity of sectoral labour demand is $-\sigma S_k$, where S_k is the share of capital in total primary factor cost. And the response of capital to the 1% wage rise is σS_l. This enables us to fill in rows 4 and 5 of Table 1 for both sectors.

Using a value-share-weighted average of the sectoral factor uses we can compute(averaging) the combined column in Table 1. Total labour demand falls by 0.095%, while total capital demand rises by the same amount.[4]

The final "naive" column shows the results we would get if we followed the usual methods of aggregating flows (adding) and elasticities (averaging). Now, S_k = S_k = 0.5, σ = 1, leading to percentage changes of ±0.5.

Thus, the factor demand response from the aggregated model is more than 5 times greater than that computed from the disaggregated model.

Table 1: How a 1% wage rise affects two industries, separately and combined.

		Sector 1	Sector 2	Combined	Naive aggregation
1	Capital value	5	95	100	100
2	Labour value	95	5	100	100
3	σ	1	1	0.19	1
4	Capital % change	0.95	0.05	0.095	0.5
5	Labour % change	-0.05	-0.95	-0.095	-0.5

As shown in the "combined" column, if we had used a σ value of 0.19 [=0.095/0.5] while using the aggregate data, we would have got the same answer as if we had added together the disaggregate results. So this example suggests that we can improve on the standard, weighted average, method of aggregating elasticities.

2 Setting up the problem: a modest ambition for computing better aggregate elasticities

We might imagine a general schema for choosing aggregated elasticities, along the lines of:

Chose σ to minimize $\Sigma_i W_i(X_i - Y_i)^2$

where σ are the aggregated elasticities, X_i results from the aggregate model, Y_i aggregated results from the disaggregate model[5], and W_i a set of weights. However, such a schema suffers from at least 2 disadvantages:

[4] You can download and run TAB files, TABLE1.TAB, TABLE2.TAB. etc, showing how each table was calculated. These are at: www.copsmodels.com/archivep.htm#tpmh0174

- For different simulations we might choose different variable results to compare and weight discrepancies differently. So the aggregated elasticities σ would be specific to a simulation.
- The schema requires that we compute results from the disaggregate model -- which is just what we wish to avoid!

Instead we propose a more modest objective: to derive a formula for CES elasticities so that own-price elasticities of demand are the same in the aggregated model as those in the aggregate model.

Most of the elasticities in most CGE models are of the CES type. Arranged in a series of nests, the CES can be made to match first-order demand responses from any CRTS production function (Perroni and Rutherford, 1995). So knowing how to aggregate CES would be a very good start.

2.1 CES demand equations in percent-change form

Because we are trying to match elasticities, it will be convenient to write the CES demand equations in percent change form.

A nest with CES σ uses inputs i=1..I with cost V_i. In percent-change form the demand equations may be written:

(1) $$x_i = z - \sigma[p_i - p_*] \qquad \text{where } p_* = \sum_k S_k p_k$$

Above
x_i = percent change in demand for input i
z = percent change in demand for the composite nest output
p_i = percent change in price of input i
p_* = percent change in average input price
S_i = cost share i = V_i/V_* where $V_* = \sum_k V_k$

Henceforth we will assume that z = 0, so that the percent change demand equation becomes:

(2) $$x_i = - \sigma[p_i - \sum_k S_k p_k]$$

The RHS implies a matrix M_{ik} of compensated price elasticities, where:

$M_{ii} = -\sigma[1- S_i]$ the own-price elasticity

$M_{ik} = +\sigma[S_k]$ the cross-price elasticity (i≠k)

3 Users and inputs

3.1 Aggregating inputs does not lead to bias

In the example of Table 1, we aggregated two different users, but the set of inputs [Capital, Labour] remained the same. Alternatively, we could aggregate the input set for a particular CES nest or user. It turns out that in this case the aggregated model, using the original CES value, gives the same answer as the disaggregate model. This is illustrated in Table 2 below.

The left hand side of Table 2 (columns 1 to 4) shows a CES=0.5 user of three inputs (Capital, Skilled Labour and Unskilled Labour) facing a 1% increase in the price of Skilled Labour. This

---

[5] We assume implicitly that we have a way to compute the Y_i. For example, if the disaggregate model results included steel output changes in separate European countries, we can compute from these (using some value weights) an aggregate European change in steel output, and compare this to the result for (steel,Europe) in the aggregated model.

leads to changes of -0.375% and 0.125% in demand for Skilled and Unskilled Labour respectively (using the own and cross- price elasticity formulae given just above). From this we can compute (column 4) the price and demand for labour as a whole.

The right hand side of Table 2 (columns 5 to 6) shows the corresponding aggregate model, where the two labour types are combined, and again the CES=0.5. It gives the same answers as we would get (Column 4) by aggregating results computed from the disaggregate model.

Table 2: Aggregation of inputs does not bias the substitution response

	σ=0.5	1 Capital	2 Skilled	3 Unskilled	4 Combined Labour	5 Aggregated: Capital	6 Aggregated: Labour
1	Value	50	25	25	50	50	50
2	% change price	0	1	0	0.5	0	0.5
3	% change quantity	0.125	-0.375	0.125	-0.125	0.125	-0.125

We can conclude from Tables 1 and 2:

- Table 1: Combining nests (or users), where the aggregate nest has a CES which is a simple average of the individual nest CES, leads to an overestimate of substitution responses. To get the 'right' answer, we need to use a smaller CES value
- Table 2: Combining inputs to a nest, using the original CES value, leads to the'right' substitution responses -- that is, the aggregated responses from the original disaggregate model.

3.2 Calculating a better combined-user CES

3.3 Example with 3 inputs

We now return to the case, described in Table 1, where several CES nests are combined into one. In Table 3 below, we have two nests/users (Sector1 and Sector2) using 3 inputs (Capital, Labour and Land). We examine the effect of combining the two sectors.

Table 3: Own-price compensated elasticities of input demand, for two industries, separately and combined.

		Sector 1	Sector 2	Combined	Implicit $\sigma = \alpha_i$	Naive aggregation
1	Capital value	40	30	70		70
2	Labour value	40	70	110		110
3	Land value	20	0	20		20
4	σ	1	1		ave=0.935	1
5	Capital own-price elasticity	-0.6	-0.7	-0.643	0.989	-0.65
6	Labour own-price elasticity	-0.6	-0.3	-0.409	0.909	-0.45
7	Land own-price elasticity	-0.8	-1.0	-0.800	0.889	-0.9

In Table 3, we have users j=1...2 using goods i=1..3 with CES σ_j and use value V_{ij} shown at top left of the table.

Let $$V_{*j} = \sum_i V_{ij}, \quad V_{i*} = \sum_j V_{ij}, \quad V_{**} = \sum_i \sum_j V_{ij}$$

Using the formula given above, the own-price elasticities are shown in rows 5-7, columns 1..2, as:

(3) $$\varepsilon_{ij} = \sigma_j[1-S_{ij}] \qquad \text{where } S_{ij} = V_{ij}/V_{*j}$$

We can average these over users to find the average disaggregate own-price elasticities:

(4) $$\varepsilon_{i*} = \sum_j H_{ij}\, \varepsilon_{ij} = \sum_j H_{ij}\, \sigma_j[1-S_{ij}] \qquad \text{where } H_{ij} = V_{ij}/V_{i*}$$

If modelled at aggregate level the aggregate own-price elasticity e_i would be:

(5) $$e_i = \sigma[1-S_i] \qquad \text{where } S_i = V_{i*}/V_{**}$$

Using the 'naive' aggregation method, $\sigma = \sum_j Q_j\, \sigma_j$ where $Q_j = V_{*j}/V_{**}$

These 'naive' own-price elasticities e_i are shown in the final column, rows 5..7.

Better estimates for the aggregate CES, α_i, might be found by equating the e_i and the ε_{i*} :

(6) $$\alpha_i[1-S_i] = \sum_j H_{ij}\, \sigma_j[1-S_{ij}]$$

(7) $$\alpha_i = [1-S_i]^{-1}\sum_j H_{ij}\, \sigma_j[1-S_{ij}]$$

These α_i are shown in the 'implicit' column, rows 5..7. If there are more than 2 inputs[6], the α_i will differ according to which input i was selected. To enforce a single CES, we take an average of the α_i :

(8) $$\sigma_{agg} = \sum_i S_i\alpha_i = \sum_i S_i[1-S_i]^{-1}\sum_j H_{ij}\, \sigma_j[1-S_{ij}]$$

The average σ is shown in the cell [row=,col=implicit] of Table 3.

4 How many elasticities?

Many models force a number of CES nests to share one elasticity. For example GTAP 6 assigns a single import/domestic CES to each commodity; although the number of such nests is:

- NCOM number of commodities
- *NUSER usually NCOM+3, [industries plus C, I, G]
- *REG number of regions

eg, for edible oil, the same import/domestic CES is used by the hotel industry in Italy and households in Peru. That is, each of the NCOM elasticities is shared by NUSER*REG nests. For fullsize GTAP, NUSER*REG might be around 8000[7].

GTAP (like other CGE models) assumes uniform elasticities because elasticity estimates (especially user- and country-specific estimates) are hard to find.

However, the uniformity assumption hinders the task of forming aggregate elasticities (USA aggregate imp/dom elasticities have been averaged using World weights!).

5 Procedure for aggregating CES elasticities

Step 1: Identify the CES nests which are to be combined.

Step 2: For each aggregate CES nest, compute the appropriate σ_{agg} value using formula (8) above.

Step 3: If several CES nests are forced (by the model) to use the same σ value, we will need to form an average of the nest-specific σ_{agg}.

---

[6] As we saw above, in the two input case $\alpha_1 = \alpha_2$. The symmetry and homogeneity restrictions on the output-constant price elasticity matrix ensure that this is so.

[7] GTAP 7 adds a region subscript, but still forces CES to be the same across users within a region.

6 Borrowing elasticities estimated using broad sectors

We have suggested above that an aggregated model should in general use smaller CES values than those used in the corresponding disaggregate model. An analogous approach might be applied to choosing elasticities for the disaggregate model if the source estimate (regression) uses quite a different level of aggregation than the model. For example, suppose an econometrician has estimated an imp/dom σ for 'food'. Suppose our model has several 'food' commodities. It would be common practice to assign the same estimated food imp/dom σ to each of the model food commodities. However, formula (8) above suggests that the disaggregate σ should in general be larger than the single estimated σ, especially if import shares vary among the several food commodities of the model.

7 Conclusion

Only a brave man would claim that most CGE elasticities were accurate to better than plus or minus 50%. Is it then silly to propose, when the database is aggregated, that we should make some small, complicated adjustment to the CES elasticities?

In defence, the current paper makes no judgement on the quality of elasticity estimates. It has a more modest aim: to devise a way (when aggregating a database to reduce solution time) of setting aggregate CES values in a way that allows the aggregated model to show similar substitution reponses to the disaggregate model.

8 References

Carlo Perroni and Thomas F. Rutherford (1995), Regular flexibility of nested CES functions, European Economic Review, Volume 39, Issue 2, Pages 335-343, https://doi.org/10.1016/0014-2921(94)00018-U.

Bruno Lanz and Thomas F. Rutherford (2016) Aggregating Compensated Price Elasticities, unpublished note.

9 Appendix

We have users j=1...2 using goods i=1..3 with CES σ_j and use value V_{ij} and:

$$V_{*j} = \sum_i V_{ij}, \quad V_{i*} = \sum_j V_{ij}, \quad V_{**} = \sum_i \sum_j V_{ij}$$

$$S_{ij} = V_{ij}/V_{*j} \quad S_i = V_{i*}/V_{**} \quad H_{ij} = V_{ij}/V_{i*}$$

When combining nests, our formula for the aggregate CES is:

$$\sigma_{agg} = \sum_i S_i \alpha_i = \sum_i S_i [1-S_i]^{-1} \sum_j H_{ij}\ \sigma_j [1-S_{ij}]$$

let $D_{ij}=S_{ij}-S_i$ (or $S_{ij}=D_{ij}+S_i$), and assume all $\sigma_j= \sigma_{disagg}$

$$\sigma_{agg} = \sigma_{disagg} \sum_i S_i [1-S_i]^{-1} \sum_j H_{ij}\ [1-S_{ij}] = K\sigma_{disagg}$$

where $$K= \sum_i S_i [1-S_i]^{-1} \sum_j H_{ij}\ [1-S_{ij}]$$

We wish to show that 0<K<1 and is smaller as the variation in cost shares across users is larger.

Let $D_{ij}=S_{ij}-S_i$ (or $S_{ij}=D_{ij}+S_i$)

then $$K= \sum_i S_i [1-S_i]^{-1} \sum_j [V_{ij}/V_{i*}]\ [1-D_{ij}-S_i]$$

using $$[V_{ij}/V_{i*}] = [V_{ij}/V_{*j}]\ [V_{*j}/V_{i*}] = [S_{ij}]\ [V_{*j}/V_{i*}] = [D_{ij}+S_i][V_{*j}/V_{i*}]$$

we get $$K= \sum_i S_i [1-S_i]^{-1} \sum_j [V_{*j}/V_{i*}][D_{ij}+S_i][1-D_{ij}-S_i]$$