

NBER WORKING PAPER SERIES

CONSISTENT TRADE POLICY AGGREGATION

James E. Anderson

Working Paper 14046

<http://www.nber.org/papers/w14046>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

May 2008

I am grateful to Will Martin for suggesting the topic of this paper to me, and for helpful comments, to G. Chris Rodrigo for assistance with preparing the data and for helpful comments. I am also grateful to the World Bank for research support. I thank the referees for their helpful comments. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2008 by James E. Anderson. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Consistent Trade Policy Aggregation
James E. Anderson
NBER Working Paper No. 14046
May 2008
JEL No. C43,D58,F10,F13,F17

ABSTRACT

Most empirical policy work requires the aggregation of policies. Trade policy aggregation exemplifies the aggregation problem poignantly, with thousands of highly dispersed trade barriers. This paper provides methods of policy aggregation that are consistent with two common objectives of empirical work. One is to preserve real income. The other is to preserve the real volume of activity in the parts of the economy being aggregated. Both objectives must be achieved for consistent multi-country policy modeling. An application to India shows that the standard atheoretic method of aggregation overstates India's real income by around 3 times the global gains from free trade.

James E. Anderson
Department of Economics
Boston College
Chestnut Hill, MA 02467
and NBER
james.anderson.1@bc.edu

Aggregation of policy variables is unavoidable in empirical work. Data limitations and computational burdens dictate that extremely detailed policies must be combined into summary indexes. The focus of this paper is on trade policy aggregation, a particularly poignant form of the problem,¹ but the discussion applies to other policy aggregation problems as well. The dirty little secret of trade policy empirical work is that atheoretic policy aggregation contaminates the results. How much could not previously be even guessed at because no consistent aggregation concept existed against which to measure standard practice.

This paper develops aggregation methods that are consistent with two common intermediate objectives of empirical work: to preserve the relationship between the policy and real income and to preserve the relationship between the policy and the real volume of activities in sectors of the economy that are relevant to the purpose of the work. An application of the methods of this paper to evaluation of Indian trade policy shows that aggregation bias is large.

“Getting the policy wedges right” is important for getting real income right, as suggested by intuition based on dead weight loss triangles: the loss is roughly proportional to the square of the policy wedge. Reasoning developed below shows that

¹ Tariffs are typically levied on some 10,000 product lines, and are discriminatory across some 200 countries. It is infeasible to build disaggregated models at the level of the 2 million categories implied by the tariff line data, particularly since information on production and consumption is only available at much higher levels of aggregation. Aggregation up to a level consistent with the available information on production and consumption allows modeling to take advantage of the detailed information on protection, while using the available information on production and consumption.

the aggregated wedges are too small by significant amounts, so welfare effects are too small in absolute value.

Getting the policy wedges right is also important for getting trade volume and domestic economic activity right. The wedges act on volume directly through the substitution effect and indirectly through the income effect. In simulation models the bias is difficult to determine *a priori* because there are offsetting effects in general equilibrium. In econometric models, the bias is unknown, since it depends on the correlation between the aggregation error and the dependent variables of the analysis. Low or negative correlation between consistent and standard aggregates is sometimes plausible, which suggests substantial bias.

The aggregation methods of this paper supply the right aggregate wedges for getting real income right and a different set of correct aggregate wedges for getting sectoral volumes right. These two distinct aggregates are required for both policy simulation models and econometric models. Using the methods, investigators can consistently aggregate tariffs or tariff equivalents of quotas from the thousands of lines found in the tariff data to the dozens of separate lines used in standard policy simulation models or the hundreds of lines (sometimes) used in econometric models of trade flows. The exposition concentrates on tariffs, noting the changes needed to treat the tariff equivalents of quotas in an appendix.

Consistent aggregation is only possible within a restrictive class of model structures. Start from a multi-good multi-country representative agent general equilibrium model with convex structure. All problems due to the heterogeneous response of

individual agents to specific policies are thus assumed away.² Even so, without separability in preferences and technology, consistent aggregation with respect to prices or quantities is impossible, as Leontief and others showed long ago. Assuming separability and homotheticity at the lower level of aggregation, two (or multi-) stage budgeting characterizes both demand and supply. Exact aggregate price indexes can then be defined that maintain consistency with exact aggregate quantities.³ These enter the simulation or econometric model as if the indexes characterized single goods. While the separable setup is highly restrictive, it includes most applied general equilibrium models. Section 1 presents the basic description of the economy.

Consistent tariff aggregation is developed starting in Section 2 with the small country case where world prices are constant. The aggregation problem with respect to real income is to aggregate tariffs on a product group (e.g., clothing) over its member product varieties (e.g., men's shirts, men's trousers) such that real income is preserved in the general equilibrium of the economy. The aggregation problem with respect to trade volume is to aggregate tariffs such that group trade volume is preserved in the general equilibrium. The procedure developed below utilizes two indexes. The trade weighted average tariff is needed to get tariff revenue right and hence to get real income right. The 'true average tariff'⁴ is needed to get the volume of activities right for given real income.

² See Blundell and Stoker (2005) for a good survey of work on the heterogeneous agents problem.

³ See for example Diewert (1981).

⁴ Anderson and Neary (2005) use this term, by analogy with the 'true cost of living index'.

The true average tariff is the tariff aggregator that consistently generates aggregate trade volume and aggregate prices in the separable product group at given real income. The concept is familiar, from index number theory, but it is not consistent with general equilibrium real income unless combined in the model with the trade weighted average tariff in exactly the right way shown below.

Bach and Martin (2001) first called attention to the problem of consistent sectoral tariff aggregation in general equilibrium. Their solution was to use a revenue constant uniform tariff along with the true average tariff (which they called the tariff aggregator for expenditure) as the basis of a compensating variation approach to consistent aggregation. While the Bach and Martin approach works for welfare analysis in a single country model, they point out that it breaks down in a global model because global payments do not balance. This paper solves the global balance problem along with the problem of preserving real income with a readily operational method.

With the methods of this paper, both the true and trade weighted sectoral tariff averages are required to consistently determine real income. In contrast, the Trade Restrictiveness Index of Anderson and Neary (2005) preserves real income in shifting from the actual vector of tariffs to a *single* uniform welfare equivalent tariff.⁵ This suggests a relationship between the welfare equivalent uniform tariff on the one hand and the pair of tariff averages used here, an issue explored below.

⁵ Their focus is mainly on an overall TRI, but they also apply it to a separable product group: the uniform tariff on all members of the group that yields in equilibrium the same real income as the differentiated tariff structure of clothing.

In a broader context, the paper is related to Goulder and Williams (2003). They propose an approximation to the measurement of excess burden in general equilibrium that improves on the standard Harberger triangle while avoiding the heavy information requirements of the theoretically correct formula. By taking account of substitution between leisure and the good subject to tax reform, much of the general equilibrium impact of the tax reform can be captured and added to the Harberger triangle. Moreover, the bias from standard practice is substantial. Similarly, this paper argues that imposing separability permits operational trade policy aggregates to be constructed that improve on the atheoretic aggregates that are commonly used, and that the bias from standard practice is substantial.

Section 3 analyzes an interdependent world economy comprised of two countries.⁶ General economic equilibrium requires market clearance, with the mechanism of adjustment being world prices. The supply side of the model is simplified to a Ricardian technology within each sector to be aggregated. With separability on the demand side there is a natural price aggregate in terms of world prices for each product group. The same tariff aggregation methods apply as in Section 2. Domestic price aggregates are wedged away from the world price aggregate by the consistent tariff aggregator. The consistent aggregator for trade volume purposes is the true average tariff. A key contribution of this paper is to provide a consistent conversion of aggregated trade volume measured in internally consistent units (utilizing the true average tariff) into aggregated trade volume measured in externally consistent units (utilizing both the true

⁶ The extension to more countries is cumbersome but the principles are the same.

average tariff and the trade weighted average tariff). This conversion is necessary for global modeling in which world markets must clear and global payments must balance.

The methods of this paper evade an existence or uniqueness problem that arises with aggregation using the TRI in the large country case. Anderson and Neary (2005, Chapter 9) and Dakhliya and Temimi (2006) note that the TRI may not exist in the large country case and where it exists it need not be unique. Essentially, the present methods, by using two different indexes, are able to get both the domestic price aggregate (and hence quantity aggregate, the trade volume) right and the tariff revenue right (and hence real income right). In contrast, the uniform welfare equivalent tariff (the TRI) is subject to the tension between the revenue (and hence real income) effects of terms of trade changes and the real income effect of higher tariffs at constant terms of trade, hence the single welfare equivalent uniform tariff may not be able to get both right in the large country case.

Section 4 presents an application to India in 1991 and 2004 using a simple computable general equilibrium model of India and the rest of the world. Tariffs at the 4 digit level for 1991 (around 1200 items) and 6 digit level for 2004 (around 5500 items) are aggregated to form two composites, final goods imports and intermediate goods imports. Aggregation bias is large, using trade weighted as opposed to consistently aggregated tariffs overstates real income for India by nearly 1%. More significantly, aggregation bias is very large relative to the gains from liberal trade. For this comparison it is necessary to control for the terms of trade effect. Simulating the effects of unilateral free trade for India at constant world prices using both trade weighted and consistent aggregation shows that the gain is understated by a factor ranging from 4 (in 1991) to 50

(in 2004). Alternatively scaled, the initial overstatement of real income is around 3 times the global gains from multilateral free trade. The bias results from the dispersion of the underlying tariffs that pushes apart the true average tariff and the trade weighed average tariff. Cutting the dispersion in half still leaves considerable bias, with the estimates converging only as the dispersion of the tariff structure disappears.

Section 5 examines the relationship between the various tariff aggregates, with emphasis on the role of dispersion in connecting the TRI (the real income equivalent aggregator) to the trade weighted average tariff. Section 6 concludes.

Appendix A examines policy aggregation with quotas. Appendix B sets out details of the application to India in 1991. An appendix available on request provides further guidance for policy modelers who may apply the methods advocated in this paper.⁷

⁷ The treatment steps through the creation of true average and trade weighted average tariffs at the sectoral level and their input into the global trade simulation model. Then consistent aggregation is illustrated with a very simple 2 country global trade model (even simpler than the India application) that has the virtue of permitting a decomposition of the effects of trade reform into efficiency effects and terms of trade effects. Here too, the real income loss due to protection using consistent aggregation is more than three times the real income loss estimated using trade weighted average sectoral tariffs.

1. The Basic Model

Trade policy affects the domestic price vectors of tradable goods (p, π) relative to a non-taxed numeraire good. The vector of goods with price vector p is to be aggregated consistently. In the background, the equilibrium prices of nontraded goods and factors are determined by market clearance as functions of the prices of traded goods and real income u , as explained further below. World prices are given by $(\bar{p}^W, \bar{\pi}^W)$ where the bar denotes that these prices are constant by the small country assumption. In Section 3 the bar is removed and world prices become endogenous. Domestic prices are wedged away from world prices by a vector of trade taxes or subsidies, $p - \bar{p}^W, \pi - \bar{\pi}^W$.

The trade expenditure function summarizes demand and supply behavior of the representative agent economy. Trade expenditure $E(p, \pi, u)$ is the difference between the expenditure function $e(p, \pi, u)$ and the gross domestic product function $g(p, \pi)$, each function having standard properties. The excess demand vectors are equal to

$E_p = e_p - g_p$ and $E_\pi = e_\pi - g_\pi$, by Shephard's and Hotelling's Lemmas. Trade tax revenue is given by $(p - \bar{p}^W)'E_p + (\pi - \bar{\pi}^W)'E_\pi$. Assume for simplicity that the only source of tax or subsidy activity comes on traded goods. No essential qualitative difference arises from this simplification: the aggregation problem remains the same when nontraded goods and factors are taxed.⁸

⁸ The policy aggregation problem for nontraded goods and factor taxes resembles that for trade policies and will have very similar solutions. It is perhaps less ubiquitous a problem because nontraded goods and factor taxes appear to be less highly differentiated

The equilibrium level of real income is determined by the international budget constraint, the value of imports at world prices cannot exceed the value of exports at world prices by more than the given international transfer b . The international budget constraint is conveniently rewritten as:

$$B(p, \pi, u; \bar{p}^W, \bar{\pi}^W) \equiv E(p, \pi, u) - (p - \bar{p}^W)' E_p - (\pi - \bar{\pi}^W)' E_\pi - b \leq 0. \quad (1.1)$$

The middle expression sets out the structure of the balance of trade function,

$B(p, \pi, u; \bar{p}^W, \bar{\pi}^W, b)$. The equilibrium condition is

$$B(p, \pi, u; \bar{p}^W, \bar{\pi}^W, b) = 0.$$

Real income u is determined by trade policies and world prices using all the information embedded in the above reduced form equation. In general, no aggregation of the prices and trade policies of the p group can do the job. Also, the determination of trade volumes $e_\pi - g_\pi$ and ‘market access’ $\bar{p}^W'(e_p - g_p)$ for given real income requires all the details of the dependence of $e_p - g_p$ on the vector p . In contrast, in the special case of separability, consistent aggregation is possible.

Specifically, the set of goods with prices p are assumed to enter preferences or production separably (weak separability is assumed) and homothetically (at the lower level of the separable preferences). Thus the price aggregator $\phi(p)$, concave and homogeneous of degree one in p , enters either as a preference aggregator ($g_\phi = 0$) or a technology aggregator ($e_\phi = 0$), with the latter representing either outputs (exports) or inputs (imported intermediates). Besides the traded goods, each country produces a

than trade taxes. The appearance may be deceptive, however, because income tax codes in many OECD countries are complex.

vector of nontraded goods with price vector h . In the background, the gross domestic product function embeds the nontraded primary factors of production.

The trade expenditure function is defined by

$$E[\phi(p), \pi, u] \equiv \max_h e[\delta\phi(p), \pi, h, u] - g[(1 - \delta)\phi(p), \pi, h],$$

where h is the price vector of nontraded goods and δ is equal to either zero or one depending on whether the tariffs being aggregated are either intermediate goods or final goods. Note that the first order condition is implied by market clearance for nontraded goods while the concavity of $e-g$ guarantees a unique global maximum. While the trade expenditure function pushes nontraded goods into the background, it remains concave and homogeneous of degree one in the traded goods prices and retains separability.

Note that this setup imposes the Armington assumption: domestically produced traded goods are not perfect substitutes for domestic products consumed at home. On top of this, it imposes two levels of separability --- between product class i and all other product classes, and between domestic and all imported varieties of product class i . To illustrate, this means that domestic peanuts (with a price that enters the vector π) are a different good from imported peanuts (with price vector p covering the varieties of peanuts imported from all the other countries in the world).⁹ The imperfect substitutes assumption is ubiquitous in empirical work because the perfect substitutes case yields predictions that are wildly far from the observable trade patterns. The extra restriction of

⁹ Manole and Martin (2006) point out that this setup is more restrictive than Armington's original assumptions where there is a single CES aggregator over domestic and imported goods.

separability is widely used in computable general equilibrium models. This separability also permits consistent aggregation.¹⁰

By assumption, the price aggregator ϕ will consistently yield the correct trade volumes for given levels of real income. Thus, by Shephard's Lemma, the trade vector for the unaggregated goods is given by $E_\pi(\phi, \pi, u)$ while the aggregate volume of trade for the p group is given by $E_\phi(\phi, \pi, u)$. The individual volumes within the p group are given by $E_\phi \phi_p$. Note, however, that the balance of trade function defined by the left hand side of (1.1) does not similarly reduce its dependence on p to the price aggregator ϕ .

2. Trade Policy Aggregation

Let T^a denote the trade weighted (on the domestic price base) average tariff:

$$T^a \equiv \sum_i w_i T_i \text{ where } T_i \equiv (p_i - p_i^w) / p_i \text{ and } w_i \equiv p_i E_{p_i} / \sum_i p_i E_{p_i} .$$

Let T^δ denote the 'true average tariff' defined by

¹⁰ It may be possible to relax the aggregation restrictions somewhat, but this must await further work. To see the issue involved, suppose that one element of p was an argument of the gross domestic product function while still appearing in the expenditure function aggregator ϕ . Try to work through the various steps taken below to achieve consistent aggregation, and notice that it is not possible. Bach and Martin show that a closed form solution is feasible for the CES case but not for their tariff revenue aggregator.

How much bias is due to imposing separability on various non-separable structures is an important issue for future work.

$$T^\delta = 1 - \phi(p^W) / \phi(p).$$

The true average tariff, applied to the world price vector, yields the same aggregate volume of trade as the actual differentiated tariff vector.

The balance of trade constraint can be consistently solved for real income in terms of (T^a, T^δ) and the other parameters of trade policy, transfers b and factor endowments:

$$E[\phi(p^W) / (1 - T^\delta), \pi, u] - T^a E_\phi \phi(p^W) / (1 - T^\delta) - (\pi - \pi^W)' E_\pi = b. \quad (2.1)$$

This equation follows from (1.1) using the implications of separability:

$$E_\phi \phi_p' (p - p^W) = E_\phi \phi(p) T^a \text{ where } T^a = \sum_i \phi_{p_i}(p) p_i T_i / \phi(p) \text{ and } \phi(p) = \phi(p^W) / (1 - T^\delta).$$

The trade volume function is given by the derivative of E with respect to ϕ , E_ϕ .

A very important subtlety is that this aggregate is in terms of quantity units that apply *internally* to the home economy. Externally, the appropriate quantity units for import volume in a world market clearance equation are in terms of world prices.¹¹ It is necessary to convert from domestic to external units by multiplying E_ϕ by

ϕ / ϕ^a where $\phi^a \equiv \phi^W / (1 - T^a)$. This follows because the world market equilibrium condition implies $-E_{\phi^W}^* \phi^W / (1 - T^a) = E_\phi \phi^W / (1 - T^\delta)$, the foreign export volume

evaluated with the average home markup must be equal to the domestic expenditure on the aggregated good. Then:

$$E_\phi \phi / \phi^a = s^\phi(\phi, \pi) e / \phi^a \text{ where} \\ s^\phi(\phi, \pi) \equiv e_\phi \phi / e.$$

¹¹ A more technical treatment helps to clarify the issue. The definition of the expenditure function in the separable homothetic case yields

$$e(p, \pi)u = \min_{x, z} \{p'x + \pi'z \mid \psi(x, z) \geq u\} = \min_{y, z} \{\phi y + \pi'z \mid \psi[f(x), z] \geq u\} = e(\phi, \pi)u.$$

Here $y = f(x)$ is the quantity aggregator for domestic purposes, dual to the price aggregator $\phi(p)$. In world markets, however, the quantities in the vector x must be aggregated with world relative prices. In the perfect substitutes case of our applications, the quantity at world prices is a simple (unweighted) sum and the world price value is the product of that sum and the single world price for the aggregate.

Notice that the aggregators (T^a, T^δ) can be calculated in a preliminary step, independently of the rest of the general equilibrium structure. This is a great advantage when some 10,000 tariffs are to be aggregated into a few dozen indexes.

Insight into why consistent aggregation requires two tariff indexes follows from the analytics of small tariff reforms. Differentiating the right hand side of (1.1) with respect to p and using the properties of the balance of trade function (for more details see Anderson and Neary, 2005) yields

$$\begin{aligned} \mu^{-1} E_u du &= (p - \bar{p}^W)' E_{pp} dp + (\pi - \bar{\pi}^W)' E_{\pi p} dp \text{ where} \\ \mu &= \left[1 - (p - \bar{p}^W)' E_{pu} / E_u - (\pi - \bar{\pi}^W)' E_{\pi u} / E_u \right]^{-1}. \end{aligned}$$

Taking account of separability (i.e., differentiating (2.1)) yields

$$\begin{aligned} \mu^{-1} e_u du &= T^a \phi E_{\phi\phi} d\phi + (\pi - \bar{\pi}^W)' E_{\pi\phi} d\phi \text{ where} \\ \mu &= \left[1 - T^a s^\phi - (\pi - \bar{\pi}^W)' E_{\pi u} / E_u \right]^{-1}. \end{aligned}$$

Notice that both the true average tariff (represented by $d\phi$ and by the ϕ argument of the share and derivative functions) and the trade weighted average tariff are required for a consistent analysis. The former is required because of its influence on the change in trade volume and the latter is required because it controls how that change in trade volume changes tariff revenue.

Relation to the TRI Aggregator

It is helpful to relate the consistent aggregation proposed above to the sectoral TRI aggregator previously proposed by Anderson and Neary. Let the TRI uniform tariff factor be given by $1 + \tau^\Delta$. This exact aggregator is implicitly defined by

$$B[\bar{p}^W (1 + \tau^\Delta), \pi, u; \bar{p}^W, \bar{\pi}^W] = B(p, \pi, u; \bar{p}^W, \bar{\pi}^W) = b. \quad (2.2)$$

See Anderson and Neary (2005) for details.

The TRI uniform tariff on the domestic price base is $T^\Delta = \tau^\Delta / (1 + \tau^\Delta)$. The uniform tariff equivalent factor on the foreign price base is given by $1 + \tau^\Delta = 1 / (1 - T^\Delta)$.

The balance of trade function in the separable case reduces to

$$B(p, \pi, u; \bar{p}^W, \bar{\pi}^W, b) = B[\phi, \pi, u; \phi(\bar{p}^W), \bar{\pi}^W, b].$$

The TRI aggregate is defined implicitly by

$$B[\phi(\bar{p}^W) / (1 - T^\Delta), \pi, u; \cdot] = B[\phi(p), \pi, u; \cdot].$$

Computation of the TRI uniform tariff for each sector to be aggregated remains computationally burdensome in that it must be calculated using the full general equilibrium structure, even as simplified by the assumption of separability. The computationally simpler alternative developed here takes further advantage of separability.

A Concrete Example

The implications of consistent aggregation are usefully illustrated by a concrete example (further developed and simulated in an appendix available on request¹²). The relationship of the TRI to the true and trade weighted average tariffs is cleanly illustrated by the example.

Suppose that the tariff ridden group to be aggregated is a final goods group while the only other tariffs are on final goods as well. The upper level expenditure function is Cobb-Douglas with parametric share parameters s^j . The lower level expenditure function is a homogeneous of degree one and concave function $\phi(p)$, such as the CES function. The budget constraint combined with the expenditure function assumption implies that:

¹² See www2.bc.edu/~anderson/AggregationAppendix.pdf.

$$e = \frac{g + b}{1 - s^\phi T^a - s^\pi T^\pi}$$

$$u = \frac{e}{\phi^{s^\phi} \prod_j \pi_j^{s^j}}$$

$$\phi = \phi^W / (1 - T^\delta).$$

The TRI uniform tariff is solved from

$$\frac{(1 - T^\delta)^{s^\phi}}{1 - s^\phi T^a - s^\pi T^\pi} = \frac{(1 - T^\Delta)^{s^\phi}}{1 - s^\phi T^\Delta - s^\pi T^\pi}. \quad (2.3)$$

Notice that both the true average tariff and the trade weighted average tariff (on the domestic price base) are needed to correctly solve the model for real income. The TRI uniform tariff is implicitly a function of these same two indexes in (2.3).

As for trade volume, e_ϕ is the volume at internal prices. For use in a global model this volume must be converted to external units, $e_\phi \frac{\phi}{\phi^a} = s^\phi e (1 - T^a) / \phi^W$ in order to

ensure that the budget constraint in terms of world prices obtains:

$e_\phi (\phi / \phi^a) \phi^W + e_\pi \pi^W = g + b$. This conversion is irrelevant for the small country, but it is crucial in a global model.

The method of aggregation extends straightforwardly to any number of separable groups to be aggregated. In terms of the Cobb-Douglas upper level example presented previously, think of each π_j as a composite price index similar to ϕ . The analyst typically imposes separability for groups such as clothing, food, consumer durables, and the like. Similarly on the intermediate goods demand structure one might impose separability on energy products, metal products and so forth. This may be less plausible than on the final demand side.

The method of aggregation also extends straightforwardly to any separable homothetic preferences at the upper level. The parametric shares such as s^ϕ become

functions $s^\phi(\phi, \pi)$ with functional forms given by such well-known structures as the CES or the translog.

Finally, the method of aggregation extends straightforwardly to tariffs on intermediate inputs, assuming that these enter the gross domestic product function separably. (Separability is a very strong restriction on the underlying technology.) Intermediate input demand is given by $-g_\phi \phi_p$. The equilibrium nominal expenditure is given by $e = [g + b - (g_\phi \phi) T^a] / (1 - s^\pi T^\pi)$ where $\phi = \phi^W / (1 - T^\delta)$ and $T^\delta = 1 - \phi^W / \phi(p)$. Real income is given by $e / \prod_i \pi^{s^i}$.

Non-separable Structures

What can be done if separability does not obtain? In that case there is no consistent aggregation. How much error is caused by the failure of consistent aggregation? This can only be determined by simulation analysis of a (set of) maintained non-separable structures relative to an erroneously imposed separable structure. By analogy with bootstrapping, draw random shocks to the maintained equilibrium system and compare the results of the true equilibrium with the approximation when separability is erroneously imposed. Separability is inescapable, but some sense of the error it entails might one day be part of reporting simulation results.

3. The Global Economy

A similar characterization of the foreign economy gives rise to functions and variables denoted with a *. Each economy distorts its domestic relative prices away from

world prices (p^w, π^w) using border taxes or subsidies on either exports or imports. There is a vector of untaxed goods with price vector r .

For simplicity, the goods to be aggregated are supplied by the foreign economy only. In scaling this assumption up for many countries and goods, the vector of exported goods to be aggregated is differentiated from the goods produced for sale in the home market and also from goods produced in the export market --- the Armington assumption.

The general equilibrium of the economies is described by a system of two balance of trade conditions and four subsystems of market clearing equations:

$$\begin{aligned}
 E(\phi, \pi, r, u) - (p - p^w)' E_\phi \phi_p - (\pi - \pi^w)' E_\pi &= 0 \\
 E^*(\phi^*, \pi^*, r, u^*) - (p^* - p^w)' E_{\phi^*}^* \phi_p^* - (\pi^* - \pi^w)' E_{\pi^*}^* &= 0 \\
 E_\phi \phi_p - g_p &= 0 \\
 E_{\phi^*}^* \phi_p^* - g_p^* &= 0 \\
 E_\pi + E_{\pi^*}^* &= 0 \\
 E_r + E_r^* &= 0.
 \end{aligned}$$

Separable structure in either final or imported input demand is imposed by the notation E_ϕ and this demand is supplied from the foreign economy only. The last two subsystems of market clearance equations do not impose this separable structure on the other taxed goods and the untaxed goods denoted by the subscript r .

Consistent aggregation in a global model requires a further restriction on supply. The simplest case arises when relative supply prices (the vector p^w) are independent of trade policies. A Ricardian technology among the goods to be aggregated is sufficient, as was the small country assumption of the preceding section. Weaker supply side separability assumptions (see below) will also permit aggregation though posing a small added computational burden.

The preceding section shows that the true average tariff is the consistent aggregate for trade volumes given real incomes while in combination with the trade weighted average tariff in the balance of trade function it provides the consistent aggregate for real incomes. The analyst needs two aggregate wedges for each country, based on which the aggregated model can be solved. The methods of the preceding section reveal how to calculate the aggregate wedges for given world prices. These wedges can be inserted into the multi-country general equilibrium model above. The solution will generate a vector of world prices, including an aggregate price for the p group.

Notice that the two balance of trade equations and the system of market clearance equations for the π goods require only the price aggregates $\phi(p^W) = \phi^W$ evaluated at world prices and the aggregate tariff wedges for each country:

$$\begin{aligned}
E[\phi^W / (1 - T^\delta), \pi, u] - T^a E_\phi \phi^W / (1 - T^\delta) - (\pi - \pi^W)' E_\pi &= b \\
E^*[\phi^{*W} / (1 - T^{*\delta}), \pi^*, u^*] - T^{*a} E_{\phi^*}^* \phi^*(p^W) / (1 - T^{*\delta}) - (\pi^* - \pi^W)' E_{\pi^*}^* &= -b. \\
E_\pi[\phi^W / (1 - T^\delta), \pi, u] + E_{\pi^*}^*[\phi^{*W} / (1 - T^{*\delta}), \pi^*, u^*] &= 0 \\
E_\phi[\cdot](1 - T^a) / (1 - T^\delta) - g_{\phi^W}^* &= 0 \\
E_{\phi^*}^*[\cdot](1 - T^{*a}) / (1 - T^{*\delta}) - g_{\phi^{*W}} &= 0 \\
E_r + E_{r^*} &= 0.
\end{aligned}$$

The system of equations determines world prices ϕ^W , ϕ^{*W} , r , r^* and π^W along with real incomes (u, u^*) as functions of the trade policy aggregators (the trade weighted average tariff and the true average tariff) for each country.

The fourth and fifth equation systems represent the implications of the separability assumption applied to the aggregates. First, each country's excess demand is met from the other country's supply only, with the notation E_ϕ denoting either final or

intermediate input demand. ϕ^W denotes the world supply price of the aggregated good sold into the home market.

Supply Side Separability

The preceding aggregation scheme allows no relative price changes on the supply side in the group being aggregated. If this is false, then in applying simulation models to trade reform, a potential inconsistency arises between the initial wedges T^δ, T^a and the equilibrium generated from the aggregated model. The initial world price vector p^{W0} used to form the wedges may not be proportional to the equilibrium price vector p^{W1} , causing shifts in the appropriate tariff aggregators.

Under a commonly used restriction in applied general equilibrium modeling, there is a simple solution. Suppose that the supply side is separable such that in the foreign economy a price aggregator $\eta(p^*)$ exists. η is increasing, convex, homogeneous of degree one and differentiable in the price vector p^* . A sufficient condition is for the goods in the aggregate to be jointly produced.¹³ The general equilibrium of production implies that the gross domestic product function is increasing and convex in η , reflecting the increasing opportunity cost of resources used to produce the joint product.

Market clearance of the home country's imported goods in the disaggregated model requires:

$$E_\phi \phi_p - g_{\phi^*} \eta_{p^*} = 0.$$

Trade reform changes elements of the tariff schedule and this causes changes in the world

¹³ Previously, the assumption was that η was a Leontief function (reflecting infinitely elasticity of transformation among the members of the group).

relative prices that clear the disaggregated markets for aggregate summary indexes $(T^a, T^{*a}, T^\delta, T^{*\delta})$ based on initial relative prices. These changes in world prices and the associated changes in disaggregated shares lead the aggregated model to be inconsistent with the ‘true’ outcomes. In principle it is possible to achieve consistency iteratively by using the market clearing sub-aggregate module above to update the shares used to calculate the tariff aggregates. Alternatively, it may be possible to develop simulation-based evidence that the share change problem is unimportant.

More generality than this means that no consistent aggregation is possible. Nevertheless, it may be possible to develop some bounds on the inaccuracy of ‘consistent aggregation’ of the type explored here.

4. Application

India in 1991 and 2004 had relatively high tariffs on average, and notably high dispersion of tariffs, especially in 1991. This assessment does not reflect India’s important liberalization of non-tariff barriers (NTBs) between 1991 and 2004 except insofar as the rise in tariffs replaces highly restrictive NTBs with less restrictive tariffs. Table 1 summarizes the picture with trade weighted average tariffs, true average tariffs and the coefficient of variation of (trade-weighted) tariffs for final goods and intermediate goods. The source data uses trade weighted average tariffs and imports for 1991 at the 4 digit CCCN level (the predecessor to the 4 digit Harmonized System classification), and at the 6 digit HS level for 2004.¹⁴ Data are taken from TRAINS with subsequent manipulation by the author. The true average tariff is calculated based on a CES aggregator with elasticity of substitution equal to 5.

| Table 1. Tariff Indexes for India, 1991 and 2004 | | | | | |
|--|------------|----|-------------|------------|----|
| Intermediate Goods | | | Final Goods | | |
| T^a | T^δ | CV | T^a | T^δ | CV |
| | | | | | |

¹⁴ This procedure suppresses tariff line dispersion at a still lower level as well as cross-country variation in tariffs due to trade preferences, presumptively biasing downward the measure of aggregation bias reported below.

| | | | | | | |
|------|------|------|-------|------|------|------|
| 2004 | .170 | .185 | .421 | .212 | .307 | .687 |
| 1991 | .122 | .218 | 1.351 | .208 | .269 | .713 |

The model used to simulate various policy experiments is that of Anderson and Neary (2005) extended to a two country world. See Appendix C for further details here. There is an aggregate export good produced by each country. On the importer’s side of the market the aggregate export is broken into many forms resulting in the 4 or 6 digit flows of final and intermediate goods imports. (There are around 600 goods in each category at the 4 digit level, with the classification followed by Anderson and Neary, 2005. There are around 2800 goods in each category at the 6 digit level.) Each country also produces a nontraded good jointly with the exported good with a Constant Elasticity of Transformation (CET) technology. The level of activity that feeds into the CET technology is produced by a CES technology that uses a single primary factor (labor) and a composite imported intermediate input. On the demand side of each economy there is a CES expenditure function that allocates expenditure between the nontraded good and the composite final import. Finally, the composite imports, both final and intermediate, are CES aggregates of the 4 digit disaggregated imports.

The parameters of the model are the various elasticities, the distribution or share parameters, the supplies of the primary factor for each country, the tariffs and the international transfer. The distribution parameters are benchmarked to 1991 or 2004 trade and national income accounts data, as are the labor supplies. The assumed values of the elasticities are presented in Table 2.

| | Demand | Joint Supply | Production |
|-------------|--------|--------------|------------|
| Upper Level | 2 | 2 | 0.7 |
| Lower Level | 5 | --- | 5 |

The rest of the world is highly simplified to a single country, Rest of World (ROW). In 1991 the ROW has a uniform 5% tariff (on the foreign price base), set up to export the amount that India imports in total (in external prices). In 2004, it was possible to construct ROW tariffs facing India from TRAINS data using the World Bank’s WITS software. All results for India turn out to be quite insensitive to the level of ROW tariffs.

All the action in the simulations comes from trade reform in India. Unilateral free trade, free trade in final goods and free trade in intermediate goods are simulated. A final simulation benchmarks the results by comparison with multilateral free trade. The chief virtue of the model is its focus on the consequences of the dispersion of tariffs, a matter that is essentially swept under the rug with the standard CGE modeling method of using trade weighted average tariffs everywhere. The results from 1991 and 2004 for India provide two examples of aggregation bias based on high dispersion tariff structures that differ significantly from each other.

India's 1991 GDP comprised about 1% of world GDP, rising to around 1.7% by 2004. Nevertheless, the structure of the model implies that no country is small; it is the only source of its exports to the rest of the world. Thus there are important terms of trade consequences of tariff reform that can offset or swamp the efficiency effects.

Table 3 presents the results of several policy simulations. The cell entries are indexes of India's real income under the various scenarios.

| Table 3. India's Real Income and Trade Policy, 1991 and 2004 | | |
|--|-------|-------|
| | 1991 | 2004 |
| <i>Initial Conditions</i> | | |
| Consistent Aggregation | .8931 | .9598 |
| Sectoral T^a Aggregation | .9008 | .9687 |
| <i>Unilateral Free Trade Experiments</i> | | |
| Free Trade | .8900 | .9353 |
| Final Trade | .8961 | .9597 |
| Intermediate Trade | .8866 | .9341 |
| Free Trade at Constant World Prices | .9009 | .9717 |
| <i>Multilateral Free Trade</i> | | |
| India's real income index | .8967 | .9436 |

The experiments reported in Table 3 provide the basis for analyzing the importance of aggregation bias relative to the gains from unilateral free trade for India. Table 4 collects several ways of scaling the aggregation bias that are illuminating. The most important conclusion to draw from the table is that the bias from using inconsistent aggregation is large.

| Table 4. Aggregation Bias for India | | |
|--|-------|-------|
| | 1991 | 2004 |
| Per cent of Real Income | 0.86% | 0.93% |
| Multiple of Gains from Free Trade, constant world prices | 0.75 | 0.98 |
| Multiple of Global Gains from Free Trade | 2.5 | 3.2 |

The first row scales the aggregation bias (the difference between the first and second rows of Table 3) by national income. For 1991, the bias is 0.86% of national income, rising in 2004 to 0.93% of real income. The next two rows of Table 4 relate the aggregation bias to the gains from free trade. In doing so, it is necessary to control in

some appropriate way for the effect of terms of trade changes. In the second row, the model is solved for equilibrium¹⁵ with the world prices of India's imports and exports frozen (the small country case). Aggregation bias ranges from 75% to 98% of the efficiency gains from unilateral free trade for India. More dramatically, the efficiency gains inconsistently measured from the trade weighted average base are only 1/4 as large as the true gains in 1991 and only 1/50 as large as the true gains in 2004. The third row nets out the terms of trade effect by imposing a utilitarian global welfare function, equal to $u + u^*$. For this comparison it is also more relevant to simulate multilateral free trade (to damp down the terms of trade effects). For 2004, the aggregation bias is 3.2 times the global gains from multilateral free trade. For 1991, the aggregation bias is 2.5 times the global gains.

The second conclusion to draw from the simulations is that terms of trade effects are important. India has higher tariffs than the rest of the world, with cuts in these tariffs generating negative terms of trade effects that are large enough to offset the positive efficiency effect that is isolated in the sixth row of Table 3. Even going to multilateral free trade does not help much, and in fact India loses in 2004 while achieving a modest gain in 1991. The size of the terms of trade effect may be too large due to the assumptions of the model. (The ROW has an elasticity of supply of exports essentially equal to elasticity of transformation due to the CET structure of the model acting in combination with the tiny share that exports to India form of the ROW production.)

The decomposition of the liberalization into final and intermediate goods tariffs in Table 3 reveals that a cut in the higher tariff, on the more elastically demanded good, will raise real income even though it has negative terms of trade effects. In contrast, unilateral free trade in intermediates cuts the lower average tariffs to zero on the inelastically demanded good. This has negative efficiency effects as well as negative terms of trade effects. These results illustrate a well-known principle of gradual trade reform. At given terms of trade, real income changes in proportion to $t'dm$ where t is the vector of specific tariffs and dm is the compensated change in the vector of imports. It is desirable

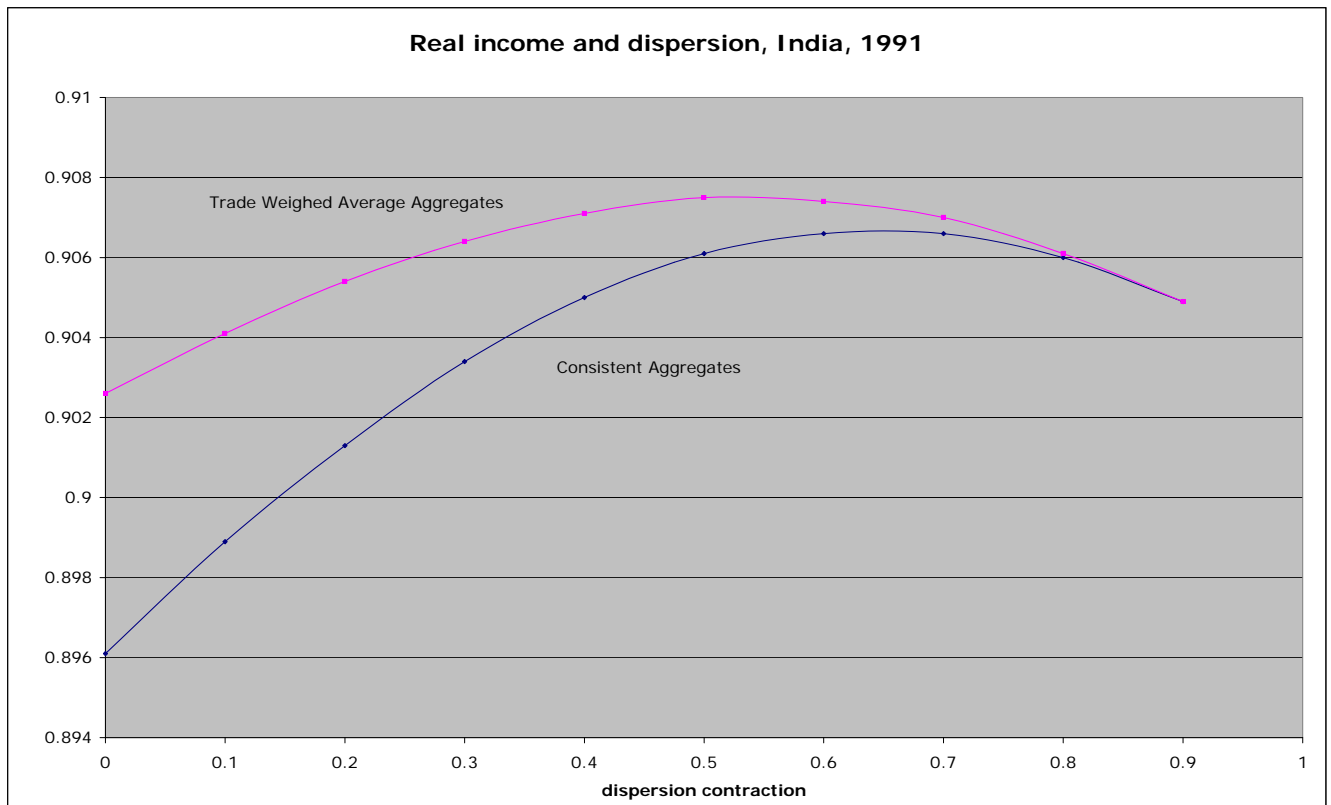
¹⁵ The solution method allows for changing prices of nontraded goods and factors as well as real income following the policy change.

to raise elements of m where the tariff is largest, and to have the largest increases in m be associated with the largest tariffs. So, to raise real income, cut high tariffs, especially where the demand is relatively elastic. Reversing this pattern will lower income.

The chart presents the real income simulation results of an arithmetic mean preserving contraction of the dispersion of tariffs, applied separately to the final and intermediate tariff schedules for India in 1991. The formula is given by

$$T^\lambda = \iota \lambda \sum_{i=1}^n T_i^0 / n + (1 - \lambda) T^0 \text{ where } \iota \text{ is the vector of ones, } \lambda \in [0,1] \text{ and the initial tariff}$$

vector is T^0 . (Because the weights in the trade weighted average tariff change, it is difficult to specify a mean preserving cut in dispersion using T^a as the mean.)



The effect of this formula on the trade weighted average tariffs is to raise the intermediate tariff (on the domestic price base) from 12.2% to 15.5% while lowering the final goods tariff from 20.8% to 11.0%. This improves real income due to the mechanism discussed above in considering the move to free trade in final and intermediate goods separately. However, there is another important mechanism at work

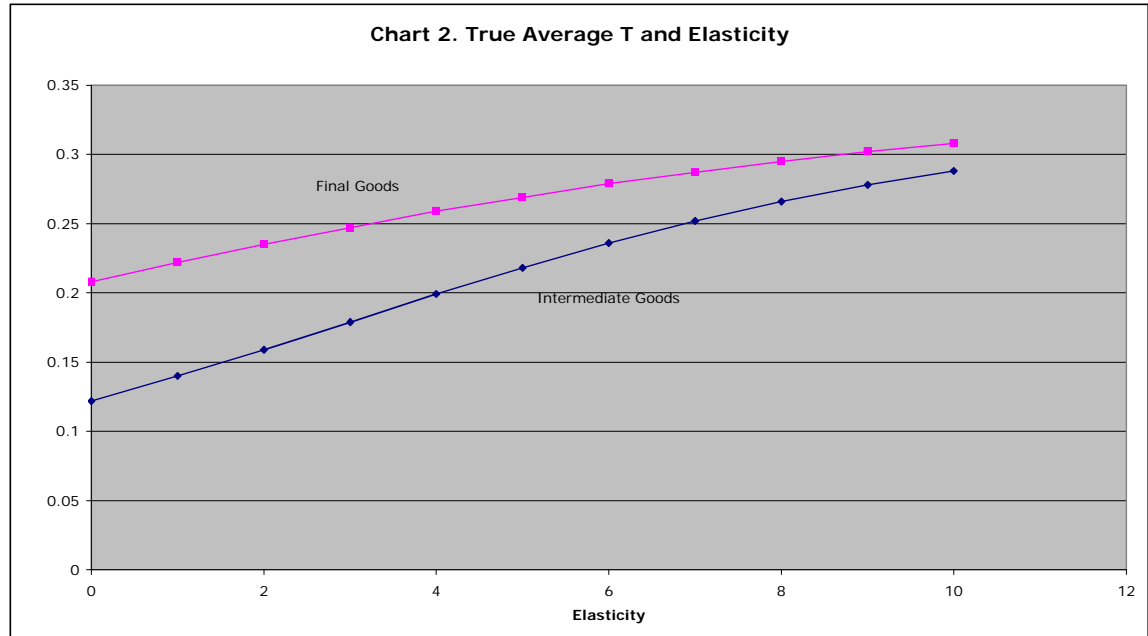
as well: the contraction of dispersion improves efficiency. The consistent aggregation method picks up this effect while the trade weighted average method misses it entirely. The gap between the two lines in the chart represents the bias due to inconsistent aggregation when the underlying trade policy includes important changes in dispersion. Essentially similar results obtain for India in 2004.

The appendix presents simulations of a different hypothetical world trade model with simpler structure and made up numbers. The simpler model eliminates nontraded goods and factors in order to more precisely separate terms of trade effects from efficiency effects and in order to see how much difference asymmetry of country size makes. Moreover, the number of tariff lines and their dispersion in the India/ROW model makes it difficult to get a feel for how dispersion drives the results.

Sensitivity Analysis

It is useful to consider the sensitivity of results to changes in a few key parameters. First of all, different assumed values of elasticities might be expected to make a big difference. Obviously the CES aggregator for import composites is sensitive to the elasticity parameter. Since the elasticity drives the difference between the true average tariff and the trade weighted average tariff, it will drive the difference in real income effects of tariff changes due to inconsistent aggregation. In contrast, the other ‘macro-economic’ elasticities have an unknown effect on the effect of inconsistent aggregation. A second assumption we test below is the sensitivity of the model to the closure involving international transfers.

The effect of the elasticity on the tariff aggregates of the 1991 data is shown the chart below. Essentially similar results obtain for 2004. The benchmark results used elasticity equal to 5. At elasticity equal to 0, the trade weighted average tariff is equal to the true average tariff.

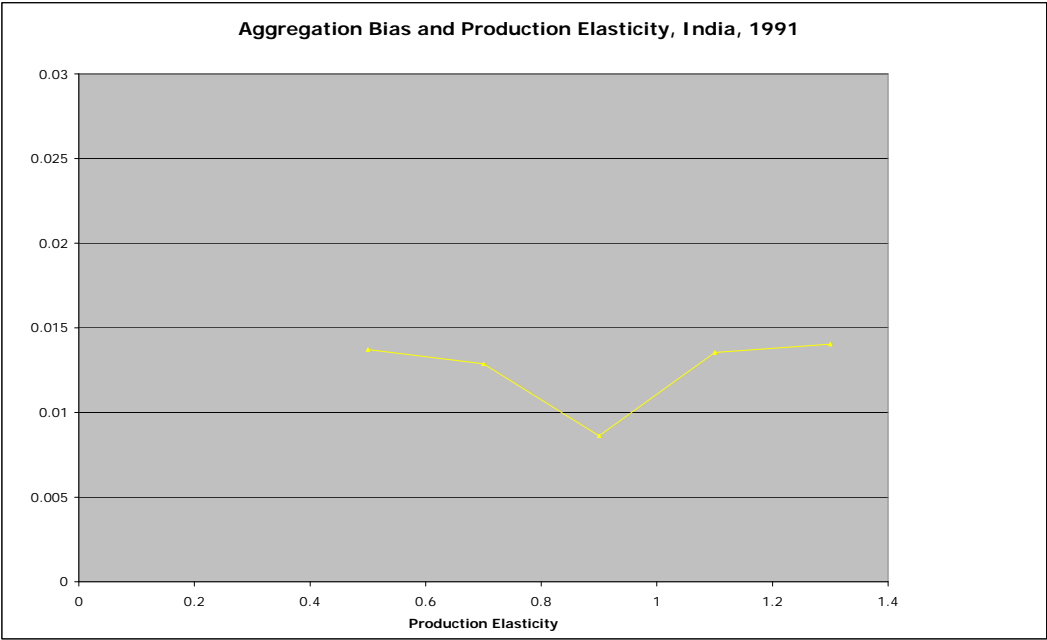
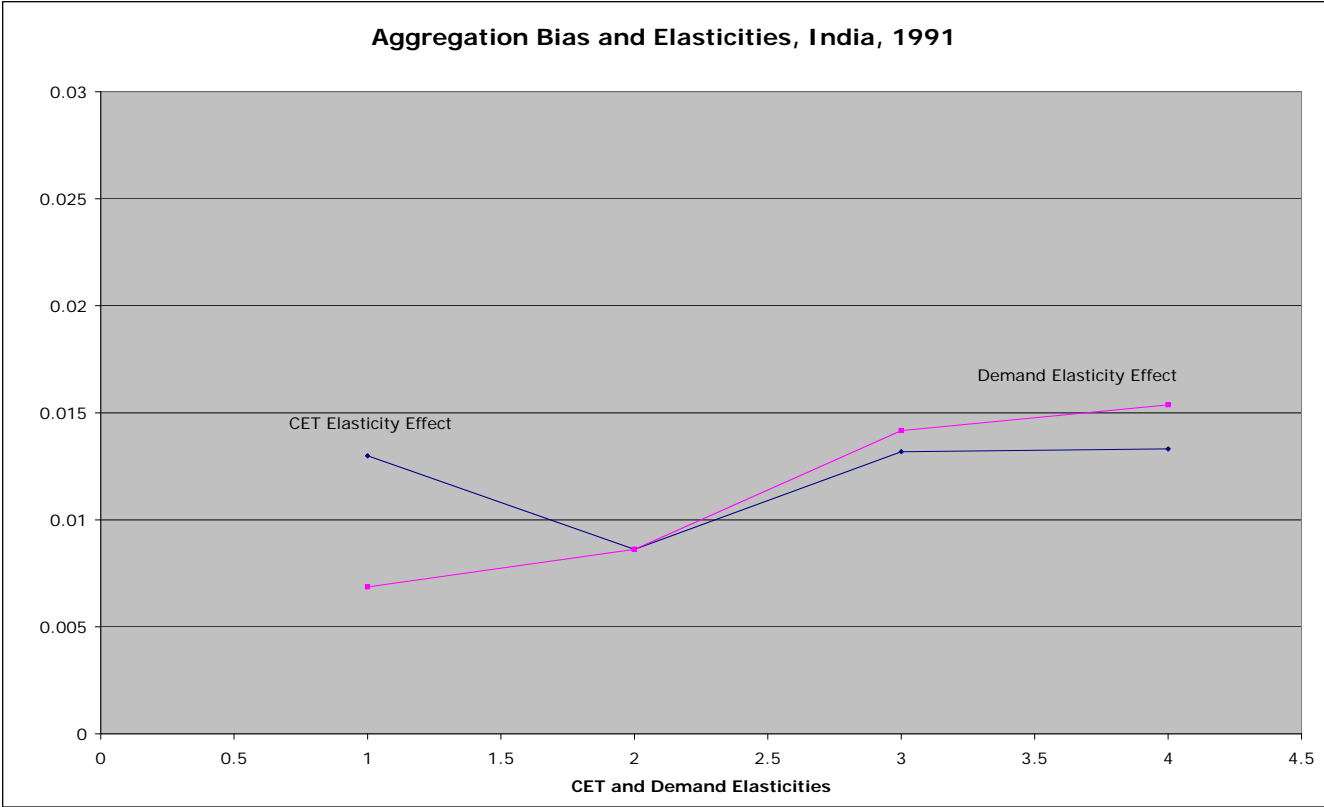


Now consider the effect on trade policy reforms of raising the elasticity to 8. (This is an elasticity commonly found in gravity model investigations of cross section trade.) Table 4 replicates the 1991 portion of Table 3 at the new elasticity value. The bias due to use of trade weighted average tariffs rises to 1.32% of national income. Notice that at the higher elasticity, unilateral free trade does pay; the efficiency effects (due basically to the difference between the trade weighted average and true average tariffs) dominate the terms of trade effect.

Table 4. India's 1991 Real Income and Trade Policy (Elasticity = 8)

| | |
|--|-------|
| Initial Conditions, Consistent Aggregation | .8891 |
| Initial Conditions, sectoral T^a aggregation | .9008 |
| Unilateral Free Trade | .8900 |
| Unilateral Free Final Trade | .8929 |
| Unilateral Intermediate Free Trade | .8859 |
| Multilateral Free Trade | .8967 |

Now consider the effect of changes in the other elasticities on the estimate of real income. The bottom line is that aggregation bias is rather insensitive to the upper level elasticities.



Finally, consider the treatment of international transfers. The benchmark model treats the apparent international transfer (the current account surplus plus an inferred net services balance) as fixed in nominal terms. The net transfer is a couple of percent of

GDP, so the magnitude is non-trivial for India. The alternative tried was to set the transfer equal to zero and adjust the apparent consumption expenditure accordingly. The effect of doing so made no qualitative difference to the results.

5. Relationship of Tariff Indexes

The application emphasizes the importance of tariff dispersion as the source of aggregation bias. Although the TRI is not required to solve global general equilibrium trade models, it is a useful diagnostic tool for analyzing the results of trade policies and their reform because it collects in a single index number the real income equivalent implications of a highly nonlinear function of the other two indexes. There are useful relationships between the true average tariff, the trade weighted average tariff and the TRI uniform tariff., along with the trade weighted variance of tariffs.

Unfortunately, the three indexes cannot be ranked except in special cases. See Anderson and Neary (2005) for a general discussion. Manole and Martin get a ranking in the CES case under an additional intuitive condition. For a self-contained treatment of the importance of consistent aggregation, this section reviews (and slightly extends) the treatment in Anderson and Neary.

From standard substitutability considerations, $T^\delta \geq T^a$. The mix of these two indexes in equation (2.1) makes it awkward to use them for analyzing the effect of changing tariff structures on welfare. In contrast, the TRI uniform tariff reduces the complexity of the tariff schedule to a single index number appearing in the balance of trade function. The TRI uniform tariff links straightforwardly to welfare. From Anderson and Neary (2005, p. 57), $T^\Delta \geq T^\delta$ if and only if tariff revenue would rise at constant real income when replacing the true average tariff with the TRI uniform tariff. There is no

general relationship possible between the TRI uniform tariff and the other two aggregates, but consideration of the structure of revenue provides more insight.

In the present model, the difference in tariff revenue evaluated with the two different uniform tariffs is given by

$$T^\delta [\bar{\phi} / (1 - T^\delta)] E_\phi [\bar{\phi} / (1 - T^\delta), \pi, u^0] - T^\Delta [\bar{\phi} / (1 - T^\Delta)] E_\phi [\bar{\phi} / (1 - T^\Delta), \pi, u^0].$$

Here $\bar{\phi} \equiv \phi(\bar{p}^w)$. The derivative of revenue with respect to a uniform tariff is signed by

$$1 + \frac{E_{\phi\phi} \bar{\phi} / (1 - T)}{E_\phi} \frac{T}{1 - T}.$$

Revenue is rising in T for small values of T , and eventually turns negative, the well known Laffer curve. Thus for small values of T , $T^\delta < T^\Delta$ while for large values of T , $T^\Delta < T^\delta$. Note that $T / (1 - T) = \tau$, the uniform tariff on the world price base. The critical value of the uniform tariff that separates the rising from the falling portion of the Laffer curve is given by $\tau^c = E_\phi / [E_{\phi\phi} \bar{\phi} / (1 - T)]$, the tariff equal to the inverse of the elasticity of demand for the product group. It is plausible that tariff schedules are usually located in the rising portion of the Laffer curve, in which case the TRI uniform tariff exceeds both of the other tariff indexes.

Tariff Averages and Variances

Anderson and Neary (2005) also show that the implications of trade reform for welfare can be summarized quite generally in terms of two tariff moments, the generalized mean and the generalized variance. The results are stated in terms of small changes but suggest that the variance of tariffs has a useful role to play in describing the welfare implications of tariff structure.

Anderson and Neary (2005) show that for the CES case:¹⁶

$$dT^\Delta / (1 - T^\Delta) = d \ln[1 / (1 - T^\Delta)] = \psi^\Delta \left[T^a dT^a + \frac{\sigma}{1 - e_\phi \phi / e} dV^a / 2 \right].$$

Here, ψ^Δ is a positive coefficient, V^a denotes the trade weighted variance of tariffs, and the differentials of the moments are evaluated using fixed weights. This expression shows that mean and variance both matter. For the most clean results, suppose that the upper level of substitutability is Cobb-Douglas. Then the square bracket term becomes

$$d \left\{ T^a \left(1 + C^2 / (1 - e_\phi \phi / e) \right)^{1/2} \right\}$$

where $C \equiv V^{1/2} / T^a$, the generalized coefficient of variation of tariffs.

Practitioners have often calculated the coefficient of variation as a supplement to average tariffs in the evaluation of trade policy. This idea was more right than theorists had previously believed. Future work may reveal that these differential expressions do tolerably well for the smaller changes that typify trade policy. In that case, empirical work for some purposes might use the formulae to approximate the behavior of the real income equivalent tariff.

6. Conclusion

Consistent policy aggregation is feasible when the group to be aggregated enters preferences or technology separably. This paper presents simple operational formulae that demonstrate the point. It also provides some simulation evidence to show that the aggregation bias problem from inconsistent aggregation is quantitatively significant. Future work should apply these methods and more thoroughly examine the biases due to inconsistent aggregation as well as some of the shortcuts suggested in the paper for simplifying computation.

¹⁶ Their equation is in terms of the TRI uniform tariff on the world price base, τ^Δ . The text expression uses the equivalence $d\tau^\Delta / (1 + \tau^\Delta) = dT^\Delta / (1 - T^\Delta)$.

References

Anderson, James E. and J. Peter Neary (2005), *Measuring the Restrictiveness of International Trade Policy*, Cambridge: MIT Press.

Bach, Christian F. and Will Martin (2001), “Would the right tariff aggregator for policy analysis please stand up?”, *Journal of Policy Modeling*, 23, 621-635.

Blundell, Richard and Thomas M. Stoker (2005), “Heterogeneity and Aggregation”, *Journal of Economic Literature*, 43, 347-91.

Cox, David and Richard Harris (1985), “Trade Liberalization and Industrial Organization: Some Estimates for Canada”, *Journal of Political Economy*, 93, 115-45.

Dakhli, Sami and Akram Temimi (2006), “An Extension of the Trade Restrictiveness Index to Large Economies”, *Review of International Economics*, 14, 678-682.

Diewert, W. Erwin (1981), “The Economic Theory of Index Numbers: a Survey”, in Angus Deaton, ed., *Essays in the Theory and Measurement of Consumer Behaviour in Honor of Sir Richard Stone*, Cambridge: Cambridge University Press, pp163-208.

Goulder, Lawrence H. and Robertson C. Williams III ((2003), “The Substantial Bias from Ignoring General Equilibrium Effects in Estimating Excess Burden, and a Practical Solution”, *Journal of Political Economy*, 111, 898-927.

Manole, Vlad and Will Martin (2006), “Keeping the Devil in the Details: A Feasible Approach to Aggregating Trade Distortions”, World Bank.

Appendix A. Quotas

This section treats the complications introduced by trade policy in the form of quotas, for simplicity returning again to the one country case.

Trade policies prominently include quotas. The most important of these, the Multi-Fibre Arrangement (MFA), was to have ended in January 2005, but pressure from the US and EU has extended its restrictions on China. Other quotas are significant in agriculture. Thus it is significant to extend consideration of consistent aggregation to cover quotas. Where tariff equivalents of individual quotas are available, the techniques of the preceding sections can be applied straightforwardly, recognizing that within group ‘tariff equivalent’ changes must be solved from the submodel that describes disaggregated trade in the quota constrained group.

Quota evaluation is usually plagued with inadequate information about tariff equivalents, a problem that appears to worsen when, as in this paper, the TRI uniform tariff is sensitive to the dispersion of the tariff equivalents. More optimism is generated by considering an important special case in which, at the margin, quota rent goes entirely to the exporter. As Anderson and Neary (1992) emphasize, only some quota rent typically accrues to the importer. The MFA for example awards the quota rights to the exporters. Tariffs on MFA-constrained categories in the US serve to retain a portion of the rent, the remainder going to exporters and importers. In the limit case, all remaining rent goes to the exporters (rationalized by their having all the bargaining power).

For each quota line, assume that there is a rent retaining tariff t^R and that the rent over and above this amount is split with ω per cent going to foreigners. The unit rent retained at home (the tariff equivalent) is given by:

$$t^R + (1 - \omega)(p - p^W - t^R) = \omega t^R + (1 - \omega)(p - p^W).$$

Now consider the evaluation of trade policy with quotas. The idea is to convert to tariff equivalents. The unit rent retained at home takes the place of the specific tariff in the balance of trade function. However, the relationship of domestic price to the ‘tariff equivalent’ and the world price is different with quotas. The virtual price (equal to the equilibrium domestic price with quotas) is given by:

$$\tilde{p} = p : E_p(p, \pi, u) = q. \quad (2.4)$$

Here, q is the quota vector. The balance of trade function with tariff equivalents of quotas is given by:

$$B(\beta\phi\pi, u; \omega, t^R) = E(\beta\phi\pi, u) - [\omega t^R + (1 - \omega)(\beta\phi - p^W)]' E_p - (\pi - \pi^W)' E_\pi - b.$$

To reduce the quota group to a single tariff, taking advantage of separability, requires treating the tariff equivalent just like the tariff vector for the p group in preceding sections, forming the trade weighted and true average tariffs on the domestic price base.

Changes in trade policy show up in changes in tariff equivalents that are not parametric. Calculation of the new tariff equivalents uses the model linking quotas and tariffs π to domestic prices, $E_p(\beta\phi\pi, u) = q$. See Anderson and Neary (1992, 2005) for details.

In the CES case, the sub-expenditure and sub-utility functions are ‘self-dual’, CES functions with the same parameters. This makes for particularly convenient versions of $E_p(\beta\phi\pi, u) = q$.

Appendix B. India and ROW CGE model

The India/ROW model is based on Anderson-Neary (2005), pp 294-303. Assume that each country produces a composite export and a non-traded good. It imports a CES composite intermediate import and a CES composite final import. The export good is effectively produced with Ricardian technology. The joint output technology between nontraded good and exports is a CET production frontier. The production of the joint output is a CES function of a single primary factor and a CES composite of imported inputs. Preferences are CES between the nontraded good and the composite final imports.

Details of the Model

The mathematical representation of the model is as follows. Let the CES price index for final goods in the home country be given by $P(p, h)$ where p is the final import goods price vector and h is the price of the nontraded final good. The equivalent variables

for the foreign country are denoted with *. Let the CES cost function for a unit of activity be given by $C(m, w)$ where m is the imported intermediate goods price vector and w is the wage rate. Let the CET activity price deflator be given by $\eta(h, r)$ where r is the price of the export good. Let z denote a unit of activity.

The constant returns to scale assumption concerning the technology implies that

$$\eta(h, r) = C(m, w). \quad (1.1)$$

Labor market clearance implies that

$$C_w z = L \quad (1.2)$$

where L is the supply of labor. Nontraded goods clearance implies that (using Shephard's Lemma)

$$P_h u = \eta_h z. \quad (1.3)$$

The balance of payments constraint implies that

$$P(p, h)u - (p - p^W)' P_p u - (m - m^W)' C_m z - wL = b. \quad (1.4)$$

Here, the W superscript denotes the world price of the goods in question. This system of 4 equations can be solved for the 4 unknowns (h, w, u, z) given the exogenous trade taxes and world prices and the exogenous supply of labor L and world transfer b .

Now consider consistent aggregation. We first of all impose that the world supply of the imported goods comes from a technology for exported goods that is Ricardian. Thus the foreign economy produces an export that has a common price for all the home economy's imports:

$$p^W = r^* \iota_p; \quad m^W = r^* \iota_m.$$

where ι denotes a vector of ones and the subscript denotes that it has the number of elements of either the vector p or the vector m . This assumption ensures that on the supply side the aggregation is consistent.

Next, we impose separability within the CES aggregators so that

$P(p, h) = P[\phi(p), h]$ where $\phi(p)$ is a lower level CES aggregator for the final goods imports. Similarly, $C(m, w) = C[\mu(m), w]$ where $\mu(m)$ is a CES aggregator for imported inputs. Then the true average tariff aggregators are defined by

$$T_p^\delta = 1 - \phi(p^W) / \phi(p)$$

$$T_m^\delta = 1 - \mu(m^W) / \mu(m).$$

Separability also lets us simplify in the balance of trade constraint to

$$Pu - T_p^a P_\phi \phi u - T_m^a C_\mu \mu z - wL = b ,$$

where the a superscript denotes the trade weighted average tariff within the group. We can solve for real income directly as

$$u = \frac{b + wL}{1 - T_p^a P_\phi \phi - T_m^a C_\mu \mu}$$

The aggregation scheme allows us to substitute for the price vectors (p, m) that have dimension in the thousands the scalar true average tariffs and the scalar world prices that have dimension 4 and the trade weighted average tariffs that have dimension 2.

Now consider closing the world model. The foreign economy is modeled in exactly parallel fashion, with all variables denoted as foreign with a *. For consistency, $b + b^* = 0$. The world markets for traded goods clear.

The subtlety with aggregation is that, while Shephard's Lemma combined with separability allows us to define natural volumes, these are in different units when the prices are different. This leads to a conversion for consistency. Markets clear according to

$$\eta_r z = P_\phi^* u^* \frac{1 - T_p^{*a}}{1 - T_p^{*\delta}} + C_\mu^* z^* \frac{1 - T_m^{*a}}{1 - T_m^{*\delta}}. \quad (1.5)$$

On the right hand side is the demand for home export goods generated by final demand and by intermediate demand, with the ratio of the tariff indexes serving to change units for consistency. On the left hand side is the supply of home export goods. An analogous expression obtains for market clearance for the export of the foreign country:

$$\eta_r^* z^* = P_\phi u \frac{1 - T_p^a}{1 - T_p^\delta} + C_\mu z \frac{1 - T_m^a}{1 - T_m^\delta}. \quad (1.6)$$

Equations (1.5)-(1.6) combine with equations (1.1)-(1.4) and the analogous equation subsystem for the foreign economy to make a 10 equation system to determine

the world prices (r, r^*) along with $(h, h^*, w, w^*, z, z^*, u, u^*)$. Homogeneity of degree zero implies that one price can be set equal to one. It is natural to normalize on the home wage rate or the foreign price of exports.

Implementation

Take the 1991 India data set utilized in Anderson and Neary (2005). Compute the ROW data for construction of ROW by taking the 1991 World Development Report totals. Adjust the export totals for ROW to match the imports of India and similarly adjust the import totals of ROW to match the export totals of India. For tariffs in ROW, assume some reasonable uniform tariffs, for example 5%.

In this model there are 4 markets to clear, ultimately. There is a nontraded good market for each country and there is a world market for the export of each country.

Details of Implementation

Data for India

GDP is the sum of payments to factors located domestically. This is wL in terms of the theory. Then $\eta z = wL + \mu C_\mu$. The second term is the sum of the intermediate imports evaluated at domestic prices. So, the left hand side can be benchmarked with data available from the WDR. The value of exports is equal to $\eta_r r$. So the value of nontraded production is equal to $wL + \mu C_\mu - \eta_r r = GDP + INTIMPORTS - EXPORTS$.

These data are used to form initial shares, which together with elasticity assumptions are used to create the CES/CET price indexes $\eta(h, r), C(\mu, w)$.

Expenditure is equal, by the balance of trade condition, to GDP+Tariff Revenue+Net Transfers. Expenditure is split between FINALIMPORTS and nontraded goods. Thus nontraded goods expenditure is equal to GDP+Tariff_Revenue+Net_Transfers-FINALIMPORTS. The nontraded expenditure share and the final imports expenditure share can be drawn from these data. Then an elasticity assumption permits us to create the true cost of living index $P(\phi, h)$.

The sub-expenditure price index ϕ and sub-cost function price index $\mu(m)$ can be created from the detailed trade flow data just as in the earlier spreadsheet model, implementing details laid out in my earlier notes.

The India data for 1991 has about 1200 tariff lines at 4 digit HS level. These are divided into roughly half intermediates and half final goods. The GDP, Exports and current account balance data are taken from the WDR. A two level CES structure is used in calculating true average tariff aggregators, with a lower level elasticity being the key parameter in aggregation.

Data for ROW

ROW GDP is obtained from WDR data. The ROW exports are set equal to India's total imports. The net transfer is set equal to the India net transfer times minus one.

The split between final and intermediate import usage in the ROW is problematic. The simulation uses the India shares.

Tariffs for ROW: Since most of the world's GDP is produced by low average tariff rich countries, a low uniform tariff of 5% is utilized. This misses important aspects of tariff structure in rich countries, but the errors here are unlikely to affect the main object of the analysis: the bias due to inconsistent aggregation.

Appendix. Illustrative Simulations

(Not for publication)

A tightly specified smaller scale model reveals the underpinnings of why consistent aggregation matters. The simulation aggregates a set of 5 tariffs that range from 3% to 150%, with the largest tariffs falling on the goods with heaviest weight in the sub-expenditure function. This tariff structure resembles apparel and agriculture protection in a number of countries. At the upper level of the model, preferences are Cobb-Douglas, with 25% of expenditure falling on the aggregated good. There is another tariff-ridden good accounting for 25% of expenditure, paying a 10% tariff. The remaining 50% of expenditure is not taxed.

All the consumption goods in the economy are imported (an extreme but inessentially simple assumption). The production of the economy is sold to foreigners, with proceeds that pay for imports. The set up is a two country general equilibrium model. The tariff-ridden traded goods (both the consumption aggregate and the other tariffed good) are perfect substitutes on the supply side. This can be rationalized for example by Ricardian technology. The tariffed and non-taxed goods are not substitutable at all on the supply side, effectively constrained by endowments.

The initial tariffs (on the world price base) and their distribution weights in the group to be aggregated are given in Table A.1. below.

| Table A.1 Initial Tariffs and Weights | |
|---------------------------------------|---------|
| Weights | Tariffs |
| 0.0149 | 0.05 |
| 0.0583 | 0.1 |
| 0.0778 | 0.03 |
| 0.3538 | 1.5 |
| 0.4953 | 0.75 |

The elasticity of substitution is set equal to 5, which is on the low side of elasticities found from gravity type models of bilateral trade. These parameters and the CES form result in the home country tariff indexes given in the first line of Table A.2.

The home and foreign country are each assumed to have an endowment of 50 for the non-taxed good and an endowment of 50 for the tariff-ridden goods. The foreign economy has a uniform 5% tariff (on the world price base) on the aggregated sector and a 10% tariff on the other tariffed sector. The model is solved for the world prices of the four goods (the non-taxed good in each country and the aggregate tariffed good on the supply side.) Homogeneity of degree zero of the world excess demand system implies that one of the prices may be set at unity.

Formally, the model is described by the following system of equations:

$$\begin{aligned}
 e &= \frac{p_T^{*W} y + p_N^{*W} z}{1 - s^\phi T^a - s^\pi T^\pi} \\
 e^* &= \frac{p_T^W y^* + p_N^W z^*}{1 - s^{*\phi} T^{*a} - s^{*\pi} T^{*\pi}} \\
 \frac{s^\phi e}{p_T^W / (1 - T^\delta)} \frac{1 - T^a}{1 - T^\delta} + \frac{s^\pi e}{p_T^W / (1 - T^\pi)} - y^* &= 0 \\
 \frac{s_N e}{p_N^W} - z^* &= 0 \\
 \frac{s^{*\phi} e^*}{p_T^{*W} / (1 - T^{*a})} + \frac{s^{*\pi} e^*}{p_T^{*W} / (1 - T^{*\pi})} - y &= 0 \\
 \frac{s_N^* e^*}{p_N^{*W}} - z &= 0.
 \end{aligned}$$

The first two equations make use of the balance of trade equation to solve for nominal expenditure. The last four equations express market clearance in the (effectively) four goods with world prices $(p_T^W, p_N^W, p_T^{*W}, p_N^{*W})$. Real incomes are obtained by deflating the equilibrium expenditures (e, e^*) by the consumer true cost of living indexes,

$[p_T^{*W} / (1 - T^\delta)]^{s^\phi} [p_T^{*W} / (1 - T^\pi)]^{s^\pi} [p_N^{*W}]^{s^N}$ for the home economy and similarly for the foreign economy.

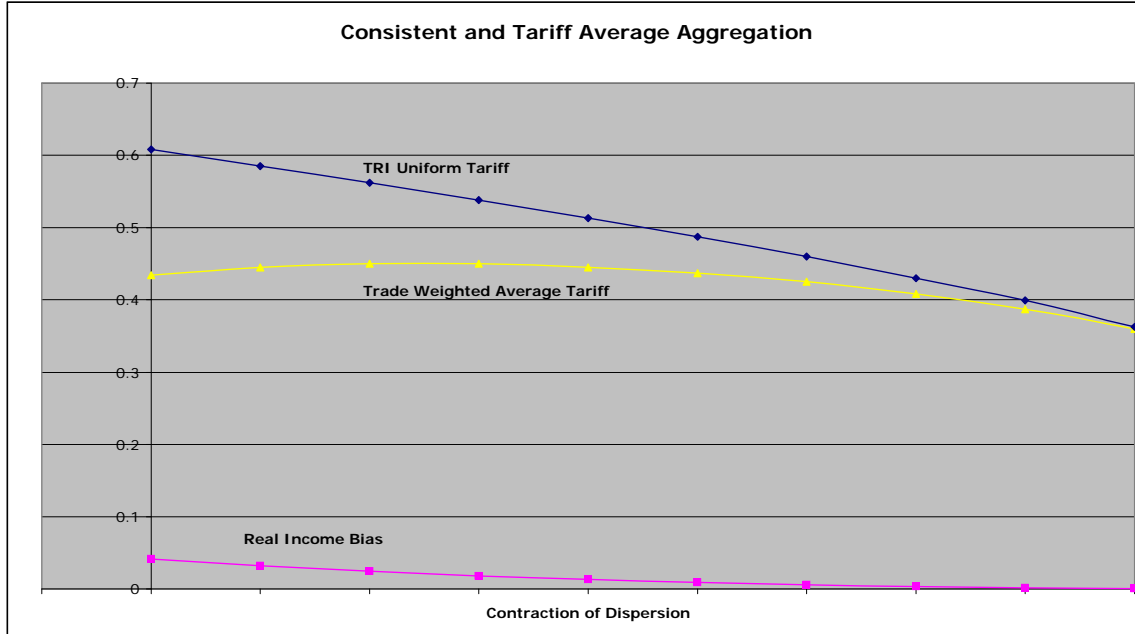
Solving for equilibrium with the parameter values given above yields the equilibrium real income listed for the home country in the second line of Table A.2. The real incomes implied by pretending that the trade weighted or the true average tariff is correct are listed in columns two and three, while column four presents the correct real income calculation.

| Table A. 2 Home Country Sectoral Tariffs and Real Income | | | |
|--|---------------------------|--------------|--------------------|
| | Trade Weighted Average | True Average | TRI Uniform Tariff |
| Tariff Index | .434 | .519 | .608 |
| Real Income | 98.62 | 97.54 | 94.69 |

These results imply a large amount of bias, amounting to over 4% of real income in the case where the trade weighted average tariff is used as if it were the right tariff (the difference between column 2 and column 4 as a proportion of column 4).¹⁷ In the present model, the welfare cost from protection amounts to 5.5%, so the bias is about 75% of the true welfare cost. Alternatively, measuring welfare costs properly more than triples the welfare cost obtained when treating the tariff using the trade weighted average. Similar bias results from using the true average tariff as if it were the proper tariff: the consistent procedure more than doubles the welfare cost estimate.

Dispersion matters a lot in driving differences between the trade weighted average tariff and the other two. Indeed, the trade weighted coefficient of variation of the tariff schedule is 0.71. The influence of dispersion on the bias due to using atheoretic average tariffs can be traced conveniently with an experiment that shrinks tariff dispersion while preserving the arithmetic mean: $T^\lambda = \lambda \sum_{i=1}^n T_i^0 / n + (1 - \lambda)T^0$ where ι is the vector of ones, $\lambda \in [0,1]$ and the initial tariff vector is T^0 . (Because the weights in the trade weighted average tariff change, it is difficult to specify a mean preserving cut in dispersion using T^a as the mean.) The chart below shows the effect of dispersion cuts of this type on the trade weighted average tariff, the TRI uniform tariff (calculated using (2.3)) and the relative bias in real income. Dispersion is decreased in 10% mean-preserving steps. On the right, all dispersion is wrung out. Even with less than half the initial dispersion, the bias is still quite significant.

¹⁷ The units in which real income is measured are unimportant; what matters is the bias, a big number by the usual standards of welfare loss from protection.



The dispersion reduction experiment lowers the TRI uniform tariff monotonically, as it does the true average tariff (not plotted), but it first raises and then lowers the trade weighted average tariff, due to the complex nonlinearity of the trade weights.

Real income rises monotonically with the arithmetic mean preserving dispersion cuts, increasing in total by around 5%. The total real income change can be decomposed into two parts: the effect of the tariff changes at constant world prices and the terms of trade effect of the change in world prices. The decomposition is based on the reduced form formula for real income:

$$u = \frac{(1 - T^\delta)^{s^\phi} (1 - T^\pi)^{s^\pi}}{1 - s^\phi T^a - s^\pi T^\pi} \frac{p_T^{*W} y + p_N^{*W} z}{(p_T^W)^{s^\phi + s^\pi} (p_N^W)^{s^N}}.$$

The first ratio on the right collects the direct tariff effects. The second term collects the terms of trade effects. The collapse of dispersion and the reduction in the trade weighted average tariff combine to worsen the terms of trade of the reforming country, costing it an additional 6.2% of real income.

The simulation results illustrate that dispersion reduction can confer a secondary benefit to a large country through a terms of trade improvement. As Anderson and Neary (2005) have shown, the trade volume (and hence the terms of trade) implications of generalized variance reduction and of generalized mean tariff reduction run in opposite directions. In the CES case, the generalized mean is the trade weighted average while the

generalized variance is the trade weighted variance divided by the share of expenditure on other goods. The decomposition of the real income changes over the interval of arithmetic-mean-preserving contractions of dispersion reveals that the terms of trade effects are at first positive and only later become significantly negative. The arithmetic-mean preserving tariff dispersion cut rule interacts with the structure of demand in this model such that the reform at first lowers trade volume and thereby improves the terms of trade. This arises because the generalized dispersion cut at first raises the generalized mean, as Chart 1 shows. Eventually the dispersion cut becomes less important than the mean cut to trade volume and the terms of trade are harmed by the trade reform.

