Measurement of Protection

James E. Anderson
Boston College and NBER

July 20, 2009

Abstract

Protection is defined broadly as government action or inaction that discriminates in favor of home producers against foreign producers. Measurement of protection has two aspects: observing or inferring protection at the product level and appropriately aggregating the highly differentiated product level protection to manageable indexes of protection. Better practice in each aspect makes a very significant difference to the measured level of protection across product groups, countries and time.

JEL Classification: F13

Prepared for the Palgrave Handbook of International Trade.

Contact information: James E. Anderson, Department of Economics, Boston College, Chestnut Hill, MA 02467, USA.

This essay surveys the measurement of protection. *Protection* is defined here to cover governmental action (or inaction) that effectively discriminates in favor of home producers against foreign producers. Protection consists of transparent formal barriers such as tariffs and quotas, less transparent formal barriers such as licensing requirements and product standards, and informal barriers such as effectively discriminatory access to law enforcement, contract enforcement and market information. The broad definition of protection is adopted because there is good evidence that informal barriers are considerably higher than formal barriers, but not uniformly so across goods or countries. A disadvantage of the broad definition is that at the informal end of the spectrum, the barriers are less obviously connected to governmental action or inaction.

Two distinct aspects of *Measurement* are treated. The first is observing or inferring protection at the disaggregated product line. The second aspect is appropriately aggregating the product line protection so observed or inferred. Aggregation is a central concern because all empirical work requires reducing the very high dimensionality of product line trade barriers[1] to a manageable size. Aggregation should be consistent with the purpose of the analysis, and inconsistent aggregation leads to highly biased results.

Section 1 deals with product line measurement of protection. The main emphasis is on inference because that is where the difficulties lie. Only a small portion of protection is directly observable as an *ad valorem* tariff set down in national tariff schedules. The problem is to go from what is observable to an *ad valorem* tariff equivalent of protection inferred.

Quotas exemplify the problem in a simple form. When no price information is available the tariff equivalent of the quota can be inferred using an economic model, for example a model of import demand and supply limited by the quantity restriction. Turning to informal protection, anecdotes of discriminatory treatment against foreign producers abound. Informal protection can be inferred from trade flows that are 'too small' to be explained within the model by observable trade costs, geography and other natural features not subject to government policy. The shortfall can be related to measures of informal institutional performance such as contract enforcement and corruption, and a tariff equivalent extracted. The gravity model has been used with success following this strategy of inference. Evidently, inference is only as accurate as the model on which it is based.

[1] For example, the tariff schedules of the US contain some 10,000 lines.

Section 2 deals with the appropriate aggregation of product line measures of protection. Until very recently, atheoretic aggregation has been the standard practice. There are several appropriate aggregators in a recently developed literature, each one ideal for a particular purpose. The survey reviews the most significant ones. Applications thus far show that atheoretic aggregation results in significantly biased aggregate measures of protection.

Section 3 discusses the complications that arise for aggregation when world prices are affected by a country's trade policies. An operational measure is proposed for protection that discriminates across countries within a product class, motivated by the discriminatory non-tariff barriers that characterize much of world trade as well as the discrimination arising from free trade areas.

1 Primary Measures of Protection

Some instruments of protection are directly observable. Here the analyst still faces a few data problems outlined below, with deepening complexity as the instruments depart from simple tax rates. Protection can be inferred from international price differences where these are available. This is usually not possible, but it is possible to infer protection from trade flows in the context of a model. See Anderson and van Wincoop (2004) for more discussion of the points developed below.

1.1 Directly Observable Measures

When countries tax trade on an *ad valorem* basis, national tariff schedules with published rates at the product line level are published and some of this data is available in online databases. There is a primary classification issue to be resolved stemming from the fact that product line definitions used by national customs authorities are not common. Since the classification issues become less severe as data is aggregated (as an extreme example, everyone's definition of imports is the same), the primary measures are often taken as atheoretic trade weighted averages at a higher level of aggregation than the finest available. The main source for this type of data is UNCTAD's TRAINS database. The World Bank's WITS software is a front end for TRAINS and the trade flow data of COMTRADE, allowing the user to construct tariff measures at the product line level for a substantial number of countries over the last 20 years.

More problematic are the many tariffs that are specific (e.g., dollars per unit). These require division by an appropriate price to convert to an *ad valorem* tariff equivalent. The difficulty is to find the appropriate price. The World Bank database has supplemented TRAINS with a significant number of conversions of specific to *ad valorem* equivalent tariffs.

1.2 Inference from Prices

Quotas can be given an *ad valorem* tariff equivalent by finding both a foreign and a domestic price for conversion using the formula $\tau = (p - p^*)/p^*$ where τ is the tariff equivalent, p is the domestic price and p^* is the foreign price inclusive of transport costs. The difficulty is to find internationally comparable price data at the product line. This is only sometimes possible.[2]

Where quota licenses are tradable, license prices can also be used to form tariff equivalents, the license price being interpreted as τ. The complications of actual trade in licenses present some difficulties in using this method. See Anderson and Neary (2005, ch. 14; 1994) for example.

Other formal or informal protection measures can similarly be given an *ad valorem* equivalent when internationally comparable prices are available.

A key problem with inference from prices is that the price differential represents *all* sources of price difference, not just those due to formal or informal protection. The apparent advantage of a model free measure is vitiated by the likelihood that the measure is contaminated by the presence of factors other than protection, such as mismeasured transport costs, costs of information, monopoly rents or mismeasured prices.

1.3 Inference from Trade Flows

Models of trade flows predict volumes conditional on supply and demand side determinants and trade frictions. 'Missing trade' suggests the work of unobservable trade frictions, and these can be related to observable variables that indicate formal or informal protection. The model structure can be used to convert the missing volume into a tariff equivalent.

The most prominent and empirically successful example of this strategy of inference is the gravity model. Because the model always gives a good fit and

[2] For an illustrative example using internationally comparable cheese prices to infer US cheese quota protection at the product line, see Anderson (1985).

exhibits stable coefficients over time and space, empirical economists have confidence that its main components capture the important determinants of the pattern of bilateral trade. Deviations from the fitted model can then be examined for indications of protection. Addtionally attractive, the gravity model has a theoretical foundation.

Gravity will be given a fuller development below, but for present purposes the end result of its derivation gives the predicted value of shipments of some generic good between country or region i and country or region j as

$$X_{ij} = \frac{Y_i E_j}{Y} \left(\frac{t_{ij}}{\Pi_i P_j} \right)^{1-\sigma}, \tag{1}$$

where X_{ij} is the predicted value of shipments from i to j at destination prices, Y_i is the total value of shipments (at destination prices) from i to all destinations, E_j is the value of spending on goods in destination j from all origins i, Y is the world aggregate of shipments $Y = \sum_i Y_i$, $t_{ij} > 1$ is the bilateral trade cost factor marking up goods shipped from i to j; and Π_i and P_j are indexes of bilateral trade costs that are defined below. σ is the elasticity of substitution parameter for the generic goods class.

The first ratio on the right hand side of (1) represents the predicted trade flow in a frictionless world. Rearranging the equation gives the frictionless prediction of trade flows as $X_{ij}/E_j = Y_i/Y$. In a frictionless world, country j's expenditure share on goods from i (on the left) is equal to the world's expenditure share on goods from i (on the right, understanding that market clearance implies $\sum_j X_{ij} = Y_i$). Behind the scenes, the assumption leading to this intuitively appealing result is that tastes are the same across the world. Trade costs act on trade through the second ratio on the right hand side of (1).[3] In effect, $t_{ij} - 1$ of the shipment melts away en route from origin to destination, so this form of trade cost is called 'iceberg-melting'.

Neither t_{ij} nor the price indexes Π_i or P_j are observable. They can be inferred econometrically by relating actual trade flow data to (1) altered on the right hand side to allow for random influences, usually represented by a multiplicative error term. A typical econometric practice is to control for the influence of Π_i or P_j with exporter and importer fixed effects. The fixed effects also encompass Y_i and E_j, which might alternatively be divided

[3] Taste differences, if admitted, would act just like trade costs on the gravity equation. Technically, gravity cannot distinguish between 'home bias' due to taste differences and home bias due to trade costs.

through into the dependent variable X_{ij}. The unobservable t_{ij} is related to observables such as direct measures of tariffs and transport costs, and proxies for trade costs such as distance, contiguity, common language and the like. Notice that t_{ij} enters (1) with the exponent $1 - \sigma$, so conversion of an estimated component of t to a tariff equivalent requires an estimate of σ. Fortunately, σ is identified with the coefficient $1-\sigma$ that attaches to directly measured trade costs such as tariffs.

It is useful for several purposes to calculate the Π and P index terms given estimates of the t's and data on the E's and Y's. The procedure will be developed in Section 3.

Consider the trade cost function that relates t_{ij} to observable proxies more carefully. t_{ij} denotes the variable trade cost factor on shipment of goods from i to j in some generic goods class. t_{ij} is usually modeled as a loglinear function of the proxies for trade costs. This is very convenient but may be false. One example is the treatment of bilateral distance. Eaton and Kortum (2002) show that allowing for shorter logs of distance to have a different effect on the log of trade than longer distances gives a significantly superior fit to the data. Another example is that insecure institutions are likely to affect some partners more than others — a common language or cultural heritage might help with contract default or extortion at the border. See Anderson and van Wincoop (2004) for more discussion.

A consequential usual practice is to ignore fixed costs of trade. Helpman, Melitz and Rubinstein (2008) set out a method to structurally identify fixed costs and their effect on volume, and report that it makes a big difference to the coefficients of the estimated t function. Their identification strategy is controversial, so many economists are skeptical of the results. (Common religion is assumed to affect the decision to enter a market, but not the volume of trade in a market already entered.) Moreover, Prusa and Besedes (2003) report that detailed product line bilateral trade flows disappear frequently, which tends to imply that fixed costs are not important.

An important limitation to gravity-based inferences about protection is that most of the literature has been highly aggregated. Typically each country produces one good (its GDP) and trades it for the GDP of all its trading partners. The theory behind the gravity model is more plausible at a disaggregated level (each country or region produces wine or clothing) and the buyers benefit from variety in their purchases of the product class. Moreover, what evidence there is from disaggregated gravity models confirms the intuition that trade costs are very different for different product classes. Differential costs over product lines would in particular be important for measurement of protection — institutional quality, membership in a FTA and nontariff barriers such as quotas would normally have different effects by product class.

1.3.1 Institutional Insecurity

Bad institutions destroy domestic trade as well as inter-regional or international trade, but it is highly plausible that foreigners are more affected, all else equal. This hypothesis has been confirmed by Anderson and Marcouiller (2002). They supplement the gravity model with national measures of institutional quality based on survey responses of business people to questions about the quality of contract enforcement and corruption. The result is that institutional quality matters a lot. Moving the quality of the Latin American countries in their study of 1996 trade flows up to the average quality of the EU would increase their trade by as much as eliminating their tariffs.

Let the institutional quality variable be denoted by Q_j for importer j. Then the component of t_{ij} that is due to insecurity is given by Q_j^γ where γ is the estimated regression coefficient. Anderson and Marcouiller find that $(1+\tau_{LA})^{1-\sigma} \approx (Q_{LA}/Q_{EU})^\gamma$, so that the tariff equivalent of the poor Latin American institutions destroys about as much trade as their tariffs on average.

It is possible that insecurity operates on the extensive margin in a way that does not lend itself so readily to a tariff equivalent measure. Crozet, Koenig and Rebeyrol (2008) model insecurity as affecting the fixed cost of exporting. Firms that are highly productive may nevertheless be excluded from exporting due the bad luck of facing big extortion demands. The resulting change in the selection can explain why, in French firm data on exports, some exporting firms serve small distant markets while not serving closer larger markets. The pure fixed cost of exports model cannot explain this pattern. On prior reasoning, however, it seems likely that extortion is quite plausibly a variable cost — either a particular shipment gets held up or not. The question of whether extortion falls mostly on variable or fixed costs is important, and in my opinion is unsettled at present.

The World Bank's *Doing Business* database has subsequently greatly enriched the set of institutional quality measures available. Unfortunately, these (and the earlier surveys) are available only as national aggregates. Presumably the quality of institutions differ in their effect on each sector.

A problem with inference about the effect of institutions on trade is that the opportunity to trade may affect the quality of institutions. Institutions change slowly, so the link to trade patterns may be weak. The reverse causality issue can be handled with good instruments for institutional quality drawn from a model of how institutions change with trade opportunities. The current state of modeling here is not yet mature enough to really help. See Dixit (2004) for some very interesting beginnings.

1.3.2 Free Trade Agreements

The best estimate of the effect of free trade agreements (FTA) on trade flows is due to Baier and Bergrstrand (2007). They report that on average a common FTA doubles trade after 10 years. An extensive earlier literature used gravity models to assess the impact of free trade agreements on trade flows with much lower estimates, often finding no significant effect. Baier and Bergstrand argue convincingly that this is due to bias caused by FTA membership being endogenously chosen by countries that already trade a lot with each other. By controlling properly for the decision to join a FTA, they quintuple the estimated effect on trade volume.

Their estimation is done with aggregate trade flows, and the FTA effect is an average effect over all members of all FTA sets. Thus there is aggregation bias over product lines and over FTA memberships, probably acting to reduce the size of estimated effects. At the product line level, an FTA will matter far more for some products than for others. Across countries, FTA's matter far more for some than for others just reasoning from the differential tariff levels they have prior to the agreement. For example, Mexico has much higher tariffs than the US or Canada, so its trade increase would be larger due to cutting a bigger bilateral barrier.

The doubling of average trade volume found by Baier and Bergstrand indicates a tariff equivalent effect of 19 percent if the elasticity is equal to 5, a tariff equivalent of 10 percent if the elasticity is equal to 8 and a tariff equivalent of 8 percent if the elasticity is equal to 10. The gravity literature places estimates of σ in the range of 5 to 10.

1.3.3 Nontariff Barriers

Measures of the presence or absence of 'non-tariff barriers to trade' have a highly significant impact on trade volume in models that use the cross-product dimension of sectoral trade flow data to identify the effect. The underlying non-tariff barrier (NTB) data are constructed from lists of potential barriers. The main source used by analysts is TRAINS, produced by UNCTAD. Some barriers are obvious, such as quotas (although even here it is possible that a quota does not bind). Others are less obvious, such as licensing requirements (which can range in effect from non-binding drivers licensing to onerous building site licensing). Technical standards and health standards have public purposes that are non-discriminatory, but are often manipulated to discriminate against imports. Antidumping cases can result in a plethora of 'remedies' that in the past have included government imposed market sharing and price fixing arrangements, but are represented in the underlying data by the presence or absence of anti-dumping actions.

In the face of this plethora of possible barriers, analysts have constructed NTB coverage ratios (the percentage of disaggregated product lines comprising the sector that have a nontariff barrier to trade) for sectors at, for example, the two digit level of aggregation. Provided that the coverage ratio has the same relationship to the volume restricted in each sector, the tariff equivalent of the NTB in sector k, τ_k, can be inferred from $NTB_k^{\beta} = (1+\tau_k)^{1-\sigma}$, where β is the estimated NTB coefficient from the cross section regression that controls for other determinants of trade flow magnitude.

Another tack is to attempt to identify elasticities econometrically from the cross country variation of prices, then use the elasticities to compute tariff equivalents of NTB's. The most notable work is by Kee, Nicita and Olarreaga (2009), drawing on their 2008 paper that estimates an import demand system yielding a set of price (own and cross) elasticities. They address the endogeneity of prices problem (due to the endogeneity of trade policy as well as violation of the small country assumption) with some success. They generate large and believable tariff equivalents.

NTB coverage data has not been used much in gravity models, reflecting the paucity of disaggregated gravity analyses. The cross country dimension of the data for a given sector identifies the volume-restricting effect of the NTB (measured by coverage ratios). As with the cross-sector regression, the identifying assumption is that the coverage ratio has a constant relationship to the volume restricted across observations, an assumption at least as dubious across countries as across commodities.

Another problem with using NTB coverage data to infer protection is posed by Trefler (1993). Political economy suggests that the presence of absence of NTB's is driven by the incentives of economic actors to lobby politicians. Trefler shows that controlling for the endogeneity of the NTB in US data greatly increased the (absolute value of the) estimate of $\beta < 0$.[4] As noted above, institutional quality is also ultimately endogenous, but over long time horizons. Thus, while the endogeneity problem also affects inference about the importance of institutions, it is presumably less significant than with NTB's.

A promising avenue of investigation is to use surveys of business persons classified by sector to obtain their rankings of the importance of various nontariff barriers to trade within their sectors. So far, the survey method has been used only at the aggregate level to measure the cross-country difficulty of doing business that is due to such factors as extortion, unreliable contractual enforcement, costly delays from officialdom.

2 Aggregation of Protection

The need for aggregation is ubiquitous in empirical work. For simple comprehensibility to the answers to such obvious questions as "How high are protectionist barriers?", reporting lists with thousand of elements is useless. Econometric inference of demand or supply relationships is seldom feasible at the level of detail of the many thousands of protectionist barriers. Applied general equilibrium simulation modeling seldom attempts to compute equilibria of models with thousands of markets. Apart from computational burden, the complexity of relationships among so many variables defeats even sophisticated readers' comprehension.

Aggregation being necessary, the key question becomes "Aggregation for what purpose?". Anderson and Neary (2005) supply answers. Each particular purpose and each particular economic environment has an ideal aggregator that preserves the relationship between the object of interest and the tariff vector being aggregated. The list of potential ideal aggregators is as long as the list of purposes supplied by empirical imagination, but the main ones detailed below suffice to illustrate a method that can be applied to extend the list.

The main focus of the literature has been on aggregation across goods. Welfare equivalent aggregation is developed first and illustrated: the uniform tariff across goods that yields the same welfare as the actual tariff vector.

[4]Tariffs, in contrast to NTB's, are constrained by WTO obligations, so they are arguably exogenous.

Then volume equivalent aggregation is developed and applied: the uniform tariff that yields the same value of trade valued at world prices. Finally, sectoral income equivalent aggregation is developed. The setting is one where world prices are taken as given — the small country case. See Anderson and Neary (2005) for treatment of ideal aggregation when world prices are endogenous to the aggregation.

The initial setting is one where world prices are taken as given — the small country case. The complexities introduced when world prices are affected by a country's tariffs are discussed in section 3.

2.1 Conceptual Base

Typical national tariff schedules contain thousands of lines. The rates are widely dispersed, by one to two orders of magnitude. A simple sense of how high are tariffs seems to require some sort of index number. Anderson and Neary (2005) describe an approach they developed which provides theoretically satisfactory yet practically implementable procedures for measuring the restrictiveness of trade policy.

The simplest context in which the aggregation problem arises is when tariffs are the only form of trade policy and the tariffs vary across products within a country. The issues are illustrated in a mainly diagrammatic analysis of an extended two-good example.

A tariff aggregator is a scalar index number that aggregates the vector of trade restrictions applying to a number of individual markets. Whether a particular index number formula is satisfactory depends on the uses to which the measure of restrictiveness is to be put. Some indices are fully satisfactory for one purpose but quite misleading for another. Other indices, lacking a clear theoretical foundation, are not satisfactory for any purpose. Three measures will be developed in some detail, each an ideal response to a different purpose.

The Trade Restrictiveness Index (TRI) is an index which aggregates trade restrictions while holding constant the level of real income. This is the natural aggregate to use in studies which attempt to link growth in income to measures of a countrys trade policy stance. It would not make sense to "explain" income growth in terms of a measure of trade policy which itself varies with income. The TRI is also the natural index to use in evaluating a countrys progress towards trade liberalization, for example, in the context of the World Banks Structural Adjustment Loans. Since loan conditionality is predicated on the assumption of a link between trade policy and income growth, it is desirable to measure the two concepts independently.

In a trade negotiations context, where foreign exporters are concerned with domestic market access, it makes sense to aggregate trade restrictions in a way which holds constant the volume of imports rather than real income. The 'volume' of imports is sensibly defined for this purpose as the value of restricted trade at external (world) prices. This leads to the Mercantilist Trade Restrictiveness Index (MTRI).

Political economy suggests strongly that protection is granted to factors of production that are specific to sectors. The pattern of earnings by job classification by sector suggests some proportion of earnings is sector specific rents. Thus it makes sense to construct an index that aggregates protection such that sector specific income is constant, sector by sector. The effective rate of protection was an early attempt to provide such an index, defined in an environment of partial equilibrium. The effective rate of protection focused exclusively on the cross effect between sectors due to intermediate stages of production: tariffs on sector i's inputs harm sectoral incomes while sector i's own output tariff helps sectoral incomes. But general equilibrium suggests other powerful cross effects: sector i's specific income is harmed by the flight of mobile factors of production to sector j induced by j's tariff even if i does not require the products of j for inputs. The sectoral income equivalent uniform tariff is defined in general equilibrium and is equal to to the effective rate of protection when general equilibrium forces become negligible.

In contrast to these theoretically based measures, it is useful to juxtapose measures which have been used in practice to aggregate across tariffs. (For simplicity the treatment here abstracts from consideration of quotas and other nontariff barriers.) These include different measures of average tariffs and alternative measures of tariff dispersion, such as the standard deviation and coefficient of variation of tariffs. The properties of these measures contrast with those of the alternative theory-based measures in a very simple context, a linear two-good partialequilibrium model. Subsequently the treatment is extended to general equilibrium.

2.1.1 Trade-Weighted Average Tariff Issues

Intuitively, tariffs should be weighted by their relative importance in some sense. The simplest and most commonly-used method of doing so is to use actual trade volumes as weights.[5] The trade-weighted average tariff is defined as:

$$\tau^a = \frac{\sum_i m_i \pi_i^* \tau_i}{\sum_i m_i \pi_i^*} \tag{2}$$

where τ_i is the ad valorem tariff rate on good i, m_i is its import volume and π_i^* its world price. This index is very easy to calculate: it equals total tariff revenue divided by the value of imports at world prices. The average tariff can be rewritten as a weighted average of tariff rates:

$$\tau^a = \sum_i \omega_i^* \tau_i \tag{3}$$

where $\omega_i^* \equiv m_i \pi_i^* / \sum_i m_i \pi_i^*$. Note that the weights ω_i^* are valued at world prices π_i^* rather than at domestic prices π_i.

Despite its convenience, the trade-weighted average tariff has well-known problems. As the tariff on any good rises, its imports fall, so the now higher tariff gets a lower weight in the index. For high tariffs this fall in the weight may be so large that the index is decreasing in the tariff rate. More subtly, tariffs have greater effects on both welfare and trade volume when they apply to imports in relatively elastic demand; but it is precisely these goods whose weights fall fastest.

Figures 2.1 and 2.2 (taken from Anderson and Neary, 2005) illustrate these considerations in a linear two-good example.

[5] An alternative is the simple (i.e., unweighted) average of tariff rates across different commodities. However, this measure has obvious disadvantages: it treats all commodities identically, and it is sensitive to changes in the classification of commodities in the tariff code.

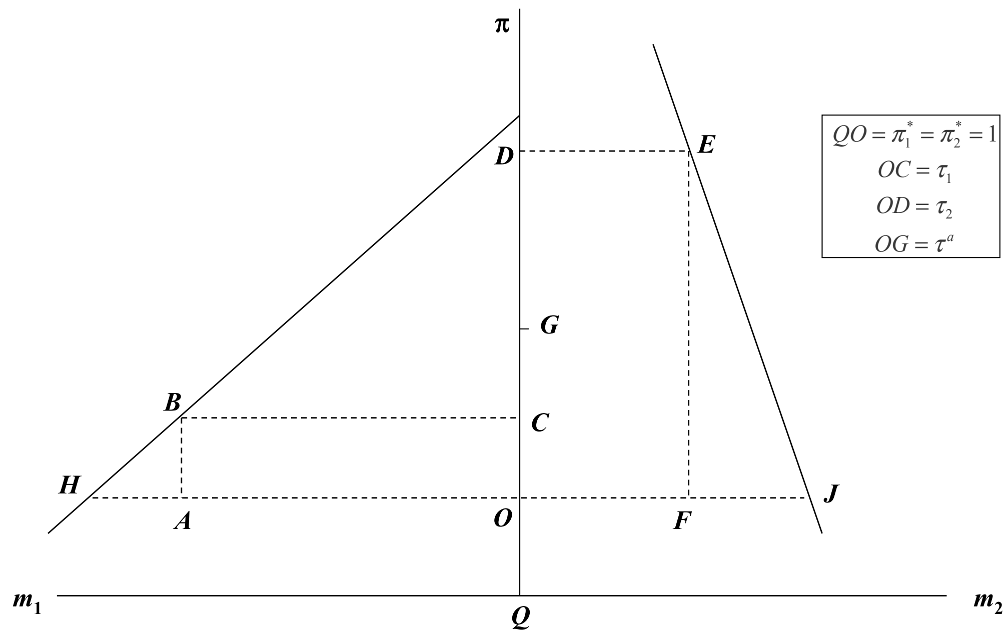


Figure 2.1: The Trade-Weighted Average Tariff:
Tariff Rates and Import Demand Elasticities Negatively Correlated

Each panel of Figure 2.1 depicts the domestic market for one of the goods, whose home import demand curve is $m_i(\pi_i), i = 1, 2$. For ease of exposition, the world prices of the two goods, π_1^* and π_2^*, are normalized at unity. Domestic producers and consumers face the tariff-inclusive prices π_i^0 represented by QC for good 1 and QD for good 2. As drawn, the import demand curve for good 1 is more elastic[6] than that for good 2, whereas good 1 has a lower tariff than good 2. So, in this example, tariff rates and import demand elasticities are negatively correlated. The trade-weighted average tariff, obtained

[6]'More elastic' is used loosely here. Strictly, good 1 has larger absolute slope: $-m_1' > -m_2'$, and the statement of the text is strictly true if changed to '...import demand schedule for good 1 is more responsive than that for good 2...'. The elasticity ranking depends on where on the demand schedules the elasticity $-m_i'\pi_i/m_i$ is evaluated. Thus the 'more elastic' statement of the text is true only for a range of tariffs and of horizontal intercepts of the linear demand schedules.

by weighting the two tariff rates by the imports (valued at world prices) of the two goods, AO and OF, is indicated by τ^a.

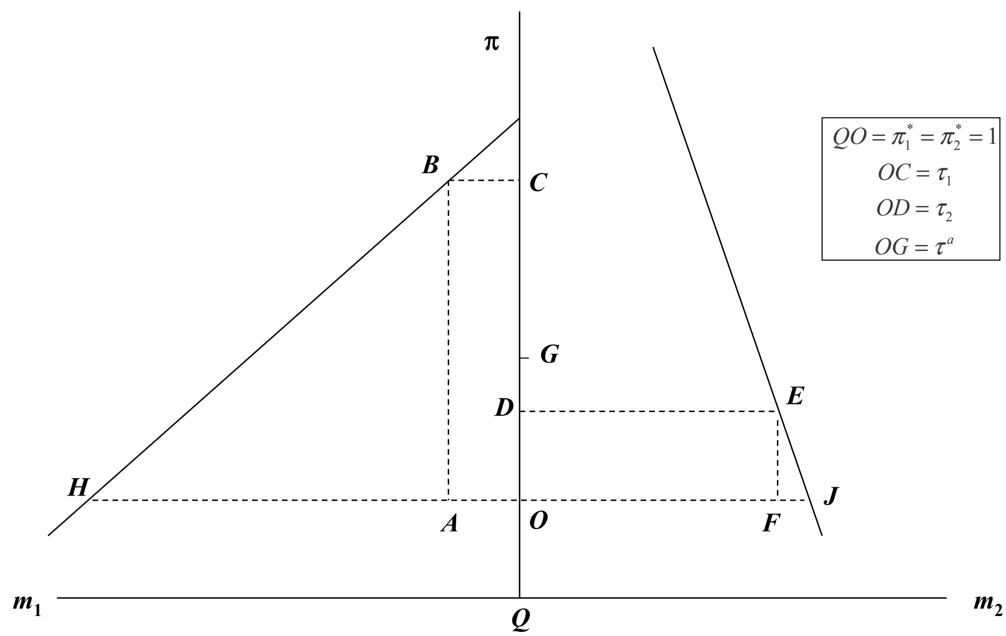


Figure 2.2: The Trade-Weighted Average Tariff : Tariff Rates and Import Demand Elasticities Positively Correlated

Next, consider a change in trade policy which leads to the situation illustrated in Figure 2.2. The two import demand functions are the same but the configuration of tariff levels is reversed: now, the correlation between demand elasticities and tariff levels is positive rather than negative. In the left-hand panel, imports of the more elastic good 1 are almost eliminated, so its high tariff receives a very low weight in the average tariff. In the right-hand panel, the low tariff on the low-elasticity good 2 receives a high weight. As a result, the calculated average tariff (again denoted by τ^a) is low, considerably lower than that in Figure 2.1. Yet, it seems intuitively obvious that trade is more restricted in Figure 2.2 than in Figure 2.1, since both welfare and the volume of trade have fallen. (Given the partial equilibrium perspective of this chapter, the deadweight loss or welfare cost of protection resulting from the two tariffs is measured by the sum of the Marshallian triangles BAH and EFJ. The volume of trade equals AF in both figures.) The index has thus moved in the wrong direction, since its value has fallen even though trade is now more restricted.

The difficulties caused by using current import volumes to construct trade-weighted average tariffs has led some authors to suggest using instead the import volumes which would prevail in free trade as weights. The choice between actual and free-trade import weights is identical in principle to that between Paasche (current-weighted) and Laspeyres (base-weighted) indices in any other branch of economics. In practice, some plausible compromise between the two (such as their geometric mean, the Fisher Ideal index) is often used.

But a central theme of the economic approach to index numbers (see, for example, Pollak (1971) and Diewert (1981)) is that the choice between alternative index-number formulae should primarily be based not on informal issues of plausibility but on the extent to which they approximate some "true" or benchmark index, which answers some well-defined economic question.

Many other weighting schemes have been proposed, but none has a superior theoretical foundation and all suffer from practical disadvantages. One possibility, discussed by Leamer (1974), is to use world exports. These have two advantages: like domestic imports, data on them are easily available; and, unlike imports, they are much less likely to be influenced by domestic tariffs. However, this virtue reflects a basic problem with using any external variables as weights: they take no account of the special features of the country being studied. Other possible sources of weights are domestic consumption or production levels. However, these also exhibit some odd features. Production shares give zero weight to tariffs on noncompeting imports; while consumption shares, like import shares, may be low for high tariffs precisely because they restrict trade so much. Finally, note that the implications of either consumption or production shares cannot be illustrated in Figures 2.1 and 2.2, since these figures as drawn are consistent with an infinite range of consumption and production levels. Thus the high tariff on good 1 in Figure 2.2 (which causes a large drop in imports and a considerable welfare cost) might get a low weight if sector 1 is less important than sector 2 in domestic consumption or production.

One implication of the previous two sections is that the problem of constructing a satisfactory aggregate tariff measure increases with the dispersion of tariff rates. This has led many practitioners to supplement weighted averages of tariff rates by measures of tariff dispersion to try and get a full picture of the restrictiveness of a tariff system. Just as with average tariffs, a key issue in choosing between different measures of tariff dispersion is which weights should be used. More generally, there is no satisfactory rule for combining the measures of average and dispersion to yield a scalar measure which might, even in principle, be comparable across countries or across time.

2.2 Welfare Equivalent Uniform Tariff

The discussion so far shows the problems with purely statistical measures such as the tradeweighted average tariff or the standard deviation of tariffs. All, in the memorable phrase of Afriat (1977), provide "answers without questions". Since they do not start from any explicit criterion of trade policy restrictiveness, their merits can be evaluated only on intuitive ad hoc grounds. And even on such grounds they do not correspond to measures of restrictiveness in any reasonable sense. A more formal approach, starting from an explicit concept of trade policy restrictiveness, is required.

The two central themes of Anderson and Neary (2005) are, first, that measures of trade policy restrictiveness should start from a formal criterion against which restrictiveness is measured; and second, that a natural criterion for an economist to adopt is the effect of the structure of trade policy on national welfare. This approach can easily be adapted to allow for other criteria, but the welfare-theoretic perspective is a natural starting point. It leads to an index number of tariffs which they call the "TRI uniform tariff" or the "welfare-equivalent uniform tariff".

It is straightforward to see how this perspective leads to an alternative measure of trade policy restrictiveness in the example given earlier. Taking welfare as the standpoint, the appropriate way of answering the question "How do we measure trade restrictiveness?" is to ask: what is the uniform tariff which, if applied to both goods, would be equivalent to the actual tariffs, in the sense of yielding the same welfare loss. Marginal deadweight loss is equal to $(\pi_i - \pi_i^*)dm_i$, the product of the change in trade volume and the difference between marginal willingness to pay π_i and marginal cost π_i^*. The change in the trade policy is responsible for the change in trade: $dm_i = m_i' d\pi_i$. The answer the question thus involves reducing dispersion in such a way that the welfare-equivalent uniform tariff is closer to the actual tariff on the high-elasticity good 1: this accords with the intuition that a high tariff on that good is more restrictive than a high tariff on good 2.[7]

The partial equilibrium approach here extends to any number of goods. To solve for the welfare-equivalent uniform tariff, write the linear import demand function for good i as:

$$m_i = \alpha_i - \beta_i \pi_i$$

where β_i is the price-responsiveness of imports of good i (i.e., the slope of the import demand curve for good i relative to the vertical axes in Figures 2.1 and 2.2). Now, recall that with linear demands the welfare loss L_i from a tariff at rate τ_i on good i equals $(\tau_i \pi_i^*)^2 \beta_i / 2$. The total welfare loss on all goods is $L = \sum_i L_i$ and so the welfare-equivalent uniform tariff τ^Δ is defined implicitly by the equation:

$$\sum_i (\tau^\Delta \pi_i^*)^2 \beta_i = \sum_i (\tau_i \pi_i^*)^2 \beta_i. \tag{4}$$

The right-hand side is the actual welfare loss from an arbitrary set of tariffs $\{\tau_i\}$; while the left- hand side is the hypothetical welfare loss from a uniform tariff rate τ^Δ. Equating the two and solving for τ^Δ gives the welfare-equivalent uniform tariff:

$$\tau^\Delta = \{\sum_i \omega_i \tau_i^2\}^{1/2},$$

where $\omega_i \equiv (\pi_i^*)^2 \beta_i / \sum_i (\pi_i^*)^2 \beta_i$. Note the differences from the formula for the trade-weighted average tariff τ^a in (2): τ^a is a weighted arithmetic mean of the tariff rates whereas τ^Δ is a weighted quadratic mean of the tariff rates; and, crucially, the weights used in constructing τ^a depend on the levels of imports, m_i, whereas those used in constructing τ^Δ depend on the marginal import responses, the β_i's. Moreover, the marginal effect of a rise in an individual tariff is invariant to the tariff in the case of average tariffs while it is proportional to the tariff in the case of τ^Δ. Marginal dead weight loss is given by $\tau_i \pi_i^* \beta_i$, so changes in τ^Δ are marginal-dead-weight-loss weighted averages of the changes in the $\{\tau_i\}$.

Now consider briefly the implications of general equilibrium for the welfare equivalent tariff. See Anderson and Neary (2005) for a thorough development; the present treatment aims to give the general idea only. (4) equates

[7] The linear demand setup of the diagram implies elasticities that vary along the demand schedule, so equating larger absolute slope with higher elasticity is only valid in a range of tariff settings.

a welfare loss due to a uniform tariff with the welfare loss of the actual tariff vector in partial equilibrium. In general equilibrium, the device analogous to the welfare loss L is the balance of trade function $B(\pi, \pi^*, u)$, where u is the real income of the representative economic agent, π is the domestic price vector of goods subject to protection and π^* is the external price vector for those goods. The welfare equivalent uniform tariff is defined by

$$B[\pi^*(1 + \tau^\Delta), \pi^*, u^0] = B(\pi^0, \pi^*, u^0). \tag{5}$$

The parallel of (5) with (4) is obvious. The same operation is being performed: a vector of differentiated tariffs is being replaced by a uniform tariff that maintains the same level of welfare. $B(\cdot)$ is an implicit function that describes the workings of an economy. Operationalizing the calculation of τ^Δ requires specifying a computable general equilibrium model of the economy. This can be demanding in terms of information, but Anderson and Neary (2005) report results for a simple and readily operational model.

B is the balance of trade deficit (the amount that must be borrowed from foreigners). The full equilibrium of the economy is expressed by the external budget constraint $B(\pi, \pi^*, u) = b$ where b is the amount borrowed. The budget constraint implies the equilibrium real income for a given domestic price vector and the assumed exogenous external price π^* and external borrowing b .The domestic price vector is driven by domestic tariff policy due to the arbitrage equation $\pi_i = \pi_i^*(1 + \tau_i)$.

The balance of trade function captures all the complex workings of the economy in a simple fashion. Behind the scenes, consumers shift their expenditures in response to trade policy changes and income changes, resources move from sector to sector, government collects and disburses tariff revenue, factor incomes rise and fall, all captured by the response of B to π. Other traded goods have prices that are constant (due to the small country assumption) while nontraded goods and factors have their prices determined endogenously in the background as functions of the price vector π and real income u. See Anderson and Neary for the many important details that go into constructing the balance of trade function and the variety of economic structures that it represents.

For present purposes it is sufficient to drive forward using B as a black box, just as auto drivers go forth with scant understanding of how their autos operate. The essential points are these. First, τ^Δ is a welfare equivalent uniform tariff. Second changes in tariffs induce changes in τ^Δ that are driven by marginal welfare responsiveness. Consider for example a change in τ_j. Differentiating (5), the required change in τ^Δ satisfies:

$$\left(\sum_i B_{\pi_i^\Delta}\pi_i^\Delta\right)\frac{d\tau^\Delta}{1+\tau^\Delta} = B_{\pi_j}\pi_j\frac{d\tau_j}{1+\tau_j}, \tag{6}$$

where $\pi_i^\Delta \equiv \pi_i^*(1+\tau^\Delta)$ and $B_{\pi_i} \equiv \partial B/\partial\pi_i$. The interpretation of B_{π_i} is the same as dL_i in partial equilibrium: the marginal cost of raising a tariff, now identified with the extra foreign borrowing that is needed to keep real income constant with the new higher tariff. B_{π_i} reduces to dL_i in the linear partial equilibrium case but is in general a more complex object that need not always be positive.

(6) shows that changes in the welfare equivalent uniform tariff are based on marginal welfare responses to changes in the actual tariffs. Just as with partial equilibrium, there is good reason to suppose that marginal welfare weights are quite different from trade weights. See Anderson and Neary (2005) for more details and evidence based on computational forms for B that τ^Δ does indeed behave significantly differently from τ^a. See Kee, Nicita and Olarreaga (2009) for recent much more detailed estimates of a partial equilibrium version of the TRI based on their estimated import demand elasticity system. See Irwin (2007) for a time series calculation of the TRI for US tariffs from 1859 to 1961.

2.3 Import Volume Equivalent Uniform Tariff

Now consider an index of tariffs which equals the uniform tariff that yields a constant volume of imports. There are two alternative valuations of imports that may be used to define 'constant volume', valuation at domestic and at foreign prices, each one being useful for a different purpose.

Valuation at domestic prices is useful in forming aggregates to be used in econometric or simulation modeling. For example, a model of the demand for imported footwear might usefully abstract from the details of types of footwear and their associated differential tariffs, giving rise to a need for a footwear tariff aggregator. Under the assumption that preferences or technology are separable with respect to the partition between footwear types on the one hand and all other goods on the other hand, an exact tariff aggregator is readily defined. Let p denote the domestic price of all other goods and let $\phi(\pi)$ be the exact price index that aggregates the individual footwear prices in the vector π. The tariff aggregator is the uniform tariff that results in the same price index value (and hence the same volume of imports). It is implicitly defined by

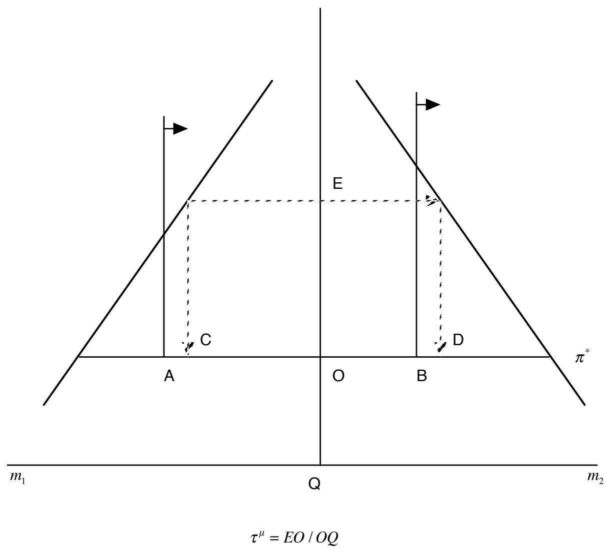
$$\phi[\pi^*(1+\tau^\delta)] = \phi(\pi). \tag{7}$$

Anderson and Neary call τ^δ the True Average Tariff because it is analogous to the true cost of living index. Because the price aggregator function ϕ is homogeneous of degree one in its arguments, (7) has an explicit solution $\tau^\delta = \phi(\pi)/\phi(\pi^*) - 1$. The marginal response of the True Average Tariff to changes in the individual tariff items is based on trade weights, but evaluated at *domestic* prices. Moreover, the exact aggregation used in τ^δ makes appropriate allowance for substitution effects among the footwear types as opposed to using the actual trade weights. See Anderson and Neary for more discussion.

The "import-volume-equivalent uniform tariff" or the "Mercantilist TRI uniform tariff" ("MTRI" for short) maintains the value of trade at world prices. This is a natural reference point for trade negotiations because foreign governments care about the value of trade at *their* prices. The Mercantilist label is used because this concept recalls the concerns of Mercantilist writers with the balance of trade.

Its behavior in the two-good example is illustrated in Figure 2.3. The initial volume $AB = CD$ is maintained in moving to the uniform volume equivalent tariff $\tau^\mu = EO/OQ$.

Figure 2.3 Volume Equivalent Uniform Tariff



The key analytic feature of Figure 2.3 is that the MTRI uniform tariff is linear in both tariff rates but increases more rapidly in the tariff on the more elastically demanded good. Moreover, the Mercantilist index behaves somewhat similarly to the welfare-equivalent uniform tariff but very differently from the ad hoc indices such as trade weighted average tariffs.

The import-volume-equivalent uniform tariff is defined implicitly by the equation:

$$\sum_i \pi_i^*[\alpha_i - \beta_i(1+\tau^\mu)\pi_i^*] = \sum_i \pi_i^*[\alpha_i - \beta_i(1+\tau_i)\pi_i^*] \tag{8}$$

The right-hand side is the total value of imports given an arbitrary set of tariffs τ_i; while the left-hand side is the value of imports which would be generated by a uniform tariff rate τ^μ. Equating the two and solving for τ^μ gives the import-volume-equivalent uniform tariff:

$$\tau^\mu = \sum_i \omega_i \tau_i$$

This has the same linear form as the trade-weighted average tariff τ^a, but the same weights $\{\omega_i\}$ as the welfare-equivalent uniform tariff, τ^Δ, in its linear partial equilibrium version.

The general equilibrium version of the MTRI follows the same basic idea of construction, but allowing for general forms of substitution in demand, supply and the link of trade policy to income. See Anderson and Neary (2005) for details. The basic building block is the real-income-compensated import demand function $m_i(\pi, p, u)$ where p is the vector of non-distorted prices (for example the world price of exported goods) and u is the representative agent's real income. The volume of trade at world prices given the real income is the compensated trade volume

$$M^c \equiv \sum_i \pi_i^* m_i(\pi, p, u). \tag{9}$$

Real income is determined by the external budget constraint

$$B(\pi, p, u) = b$$

resulting in equilibrium real income

$$u = U(\pi, p, b). \tag{10}$$

Substituting (10) into (9) yields the trade volume function

$$M(\pi, p, b) \equiv M^c[\pi, p, U(\pi, p, b)]. \tag{11}$$

The volume equivalent uniform tariff is implicitly defined by

$$M(\pi^*(1+\tau^\mu), p, b) = M(\pi, p, b). \tag{12}$$

A sense of how the tariff aggregator τ^μ is related to its components is obtained by differentiating (12) with respect to π_j and solving for the resulting change in τ^μ:

$$\left(\sum_i M_{\pi_i} \pi_i^\mu\right) \frac{d\tau^\mu}{1+\tau^\mu} = M_{\pi_j} \pi_j \frac{d\tau_j}{1+\tau_j}. \tag{13}$$

Here, $\pi_i^\mu \equiv \pi_i^*(1+\tau_i)$ and $M_{\pi_i} = \partial M/\partial \pi_i$.

The responsiveness of the MTRI uniform tariff to changes in its component tariffs is given by marginal volume weights $M_{\pi_j}\pi_j$. These weights incorporate the substitution effects on the demand and supply sides of the economy but also the income effect of price changes acting through the factor markets and through the distribution of tariff revenue. Computation of simulation models is required to actually calculate τ^μ but its logic follows the partial equilibrium version. It is clear that the weights differ from those of the trade weighted average tariff.

Computations presented in Anderson and Neary (2005) suggest that in practice the MTRI uniform tariff has quite different implications than the trade weighted average tariff. Kee, Nicita and Olarreaga (2009) present much more detailed estimates of a partial equilibrium version of the MTRI. Effectively this is a compensated MTRI based on (9):

$$\tau^{c\mu} : M^c(\pi^*(1+\tau^{c\mu}), p, u) = M^c(\pi, p, u).$$

2.4 Sectoral Income Equivalent

Tariff structure is often important to specific factor owners. For example, the US auto industry opposed the Bush administration steel tariffs in 2002 because the tariffs would raise the price of an important input into auto production. Of course, auto makers benefit from tariffs on imported motor vehicles at the same time (and from a differentiated tariff structure on different motor vehicle types) and lose from tariffs on other inputs such as textiles. These observations suggest the usefulness of an aggregate measure of the protection given by the US tariff structure to its auto industry. The suggestion led to the development of the *effective rate of protection* (Anderson, 1998): the uniform tariff on inputs and outputs that yields the same sector specific factor income as the actual tariff structure. In the special case of partial equilibrium with fixed input coefficients and fixed output coefficients (or a single output), the effective rate of protection has a very simple form (Corden, 1965). The simple form is likely to be seriously misleading (Anderson, 1998, presents a computational example where this is so) but its failings have not prevented it from being widely used in applied trade policy analysis.

To sketch out the main idea, think for example of sector specific labor skills such as auto workers possess, earning them a premium over what they could obtain in their next best alternative job. Let the earnings to the sector specific factor in sector j be denoted as ρ_j. With competitive production and the assumption of cost minimizing behavior, ρ_j is the 'restricted profit function'. Let the price vector for nontraded inputs be denoted w while the domestic price vector for traded outputs and inputs is given by π. In partial equilibrium, for given w, the effective rate of protection τ^e is defined implicitly as

$$\rho_j[\pi^*(1+\tau^e), w^0] = \rho_j(\pi^0, w^0). \tag{14}$$

The properties of the restricted profit function can be used to derive the 'marginal profit weights' that arise from differentiating (14), following the lines of the preceding analysis of welfare equivalent and volume equivalent tariffs.

In the special case where there is only one output and where the input-output coefficients are fixed, (14) gives an explicit solution for τ_j^e as

$$\tau_j^e = \frac{\tau_j - \sum_i \alpha_{ij}\tau_i}{1 - \sum_i \alpha_{ij}}, \tag{15}$$

where the α's are the intermediate input cost shares at external prices. This is the formula usually used for the effective rate of protection. It is easy to verify that if all tariffs on the right hand side are in fact equal to τ_j^e, then the equation is satisfied. Moreover, the formula implies the common sense implication that escalation in tariff structures affords extra protection to sectors:

$$\tau_j^e = \tau_j + \frac{\sum_i \alpha_{ij}(\tau_j - \tau_i)}{1 - \sum_i \alpha_{ij}} > \tau_j$$

when $\tau_j > \sum_i \omega_{ij}\tau_i$ where $\omega_{ij} \equiv \alpha_{ij}/\sum_i \alpha_{ij}$. This condition says that the output tariff exceeds a weighted average of input tariffs where the weights are the traded input share weights.

The difficulties with the usual form of the effective rate of protection are two. First, the assumption of fixed input and output coefficients is not realistic; substitutability of input mixes and output mixes is significant even in the short run. In partial equilibrium it is possible to specify a form of the restricted profit function and use its properties to derive an exact effective rate of protection as in (14). This exact effective rate of protection has the same relation to (15) as the True Average Tariff τ^δ has to the trade weighted average tariff τ^a.

More importantly, both of the preceding definitions of the effective rate of protection are defined in partial equilibrium, holding constant w^0. Constant factor prices is never a plausible assumption because the operation in (14) involves changing the entire tariff structure of the economy. In general equilibrium the factor prices w are determined by factor market clearance as a reduced form function of the trade policy that drives the demand for factors of production. Formally, $\partial\rho_j/\partial w$ gives the vector of demands for intersectorally mobile factors due to the cost-minimizing property of the restricted profit function. Market clearance for the mobile factors is given by $\sum_j \partial\rho_j/\partial w = v$ where v is the vector of mobile factor endowments. The market clearance conditions solve for the equilibrium factor prices w as functions $W(\pi, v)$ of prices π and the factor endowments. The general equilibrium analog to ρ_j is $R_j(\pi, v) \equiv \rho_j[\pi, W(\pi, v)]$. Using the general equilibrium analog to ρ_j and the same form of definition of τ_j^e as in (14) leads to a general equilibrium version of effective rate of protection.[8] See Anderson (1998) for details. Computations reported there show that the proper general equilibrium effective rate of protection has low correlation with effective protection calculated by the usual method.

3 Multi-country Issues

Measurement and aggregation of protection in a multi-country world presents two major issues. First, trade policies discriminate across countries. What is the appropriate protection aggregator across trading partners? Second, because the world market price is affected by tariffs, the incidence of the tariff, the proportion of the tax that falls on the seller's price or the buyer's price, becomes an key endogenous variable.

Discriminatory policy is discouraged under the World Trade Organization (WTO) rules (the Most Favored Nation principle), but with important exceptions. Free trade agreements are an explicit exception to the MFN principle. Their prominence means that discriminatory tariffs to affect substantial portions of world trade. There is also tacit acceptance by WTO members of many nontariff barriers (such as the 'voluntary export restraints' currently affecting US and EU imports from China in textiles and apparel).

[8]More complex economic structures are readily encompassed by essentially the same modeling strategy, encompassing complications such as nontraded goods, economies of scale or monopolistic competition.

Finally, there is explicit acceptance of restrictions (including market share arrangements, price fixing schemes and voluntary export restraints as well as tariffs) arising from anti-dumping policies so long as a country's anti-dumping regulations satisfy the WTO code.

In principle, the methods of Section 2 using fixed world prices could apply to discriminatory policies, treating as separate goods the products imported from each country. For example, within a product class such as wine (or cabernet sauvignon), national varieties differ in their appeal to consumers. Despite EU and US border policies that discriminates between French wine and California wine, EU and US consumers purchase both. Given adequate detail on wine tastes, the welfare equivalent and volume equivalent TRI's can be constructed. Another dimension of differentiated protection takes the point of view of the exporter. The exporter's goods face different treatment across trading partners. For example, developing country exports (such as apparel) typically face high protection from developed countries whereas developed countries typically have low tariffs on each other's exports (such as high quality clothing, or autos). Kee, Nicita and Olarreaga (2009) offer a partial equilibrium export volume equivalent index to capture 'average' restrictiveness from the exporter's viewpoint, the Market Access-Overall Trade Restrictiveness Index (MA-OTRI). The MA-OTRI gives an answer to the question, what uniform tariff levied by trade partners, across goods and partners, would yield the same volume of exports as the current tariff structure? Kee *et. al* find that indeed poorer countries do face higher barriers to their exports on average.

Unfortunately, the price-taking assumption of Section 2 and of Kee *et al.* becomes dubious or even untenable when measuring policy aggregates that discriminate among finely differentiated goods. California wine sellers face downward sloping demand schedules in France while French wine sellers face downward sloping demand schedules in California. At a formal descriptive level, Anderson and Neary (2005) treat the calculation of the TRI and MTRI when world prices are endogenous to the aggregation, but operationalizing their method requires computational models of substantial detail and complexity. A common approach to computational models makes extensive use of separability to aggregate tariffs within product classes using the True Average Tariff (7). This method at least guarantees the correct relationship between aggregate trade volume and the True Average Tariff, but it fails to properly connect tariff revenue and the tariffs. Anderson (2009) proposes an operational solution that uses both the True Average Tariff and the trade weighted average tariff to get both volume and revenue right in the multi-country setting with endogenous world prices. Unfortunately, all applied general equilibrium models are seen by most economists as suspect because they do a poor job of predicting trade outside the benchmark year.

A promising alternative is to focus on a single sector at a time and deal with appropriate aggregation of policy across countries in the context of the gravity model, drawing credibility from its generally good empirical fit. The analysis below sets out a readily operational uniform border policy index that consistently aggregates discriminatory policies across countries within a given product line. The idea is to replace the set of discriminatory import policies with a set of uniform tariff equivalents that preserve aggregate trade volumes in all the trading partners. This solution turns out to preserve world prices, even though world prices are generally endogenous in the model. (Thus the supporting general equilibrium between sectors of each country in the world economy is not disturbed.)

Section 3.1 sets out the structural gravity model to provide the base of analysis. Section 3.2 derives volume equivalent aggregators for discriminatory protection structures within a product class, one from the point of view of the importer and the other from the point of view of the exporter.

3.1 Structural Gravity

Begin with definitions of variables for some generic class of goods. Let i denote a country of origin and let j denote a country of destination. Let X_{ij} denote the value of shipments at destination prices from i to j. Further, let E_j denote the expenditure at destination j on goods from all origins, while Y_i denotes the sales of goods at destination prices from i to all destinations. Expenditure levels, the E's, and sales levels, the Y's, are determined in an upper level general equilibrium allocation that is exogenous for present purposes. The budget constraints (one for each country's total expenditure on each goods class) and the market clearance equations (one for each goods class for goods from each country of origin) together with a CES demand specification combine to yield the gravity model.

The CES demand function (for either final or intermediate products) with competitive pricing[9] gives expenditure on goods shipped from origin i

[9]The extension to monopolistic competition is an inessential complication because the markups in this setting are constant.

to destination j as:

$$X_{ij} = (\beta_i p_i^* t_{ij}/P_j)^{1-\sigma} E_j. \tag{16}$$

Here, the value of shipments includes the trade costs while p_i^* is the factory gate price and β_i is a CES share parameter. The price index is $P_j = [\sum_i (\beta_i p_i^* t_{ij})^{1-\sigma}]^{1/(1-\sigma)}$. To see this, sum (16) and use the budget constraint $\sum_i X_{ij} = E_j$ to simplify.

Next, impose market clearance:

$$Y_i = \sum_j (\beta_i p_i^*)^{1-\sigma} (t_{ij}/P_j)^{1-\sigma} E_j. \tag{17}$$

Define $Y \equiv \sum_i Y_i$. In a world with globally common CES preferences, the expenditure shares must effectively be generated by

$$(\beta_i p_i^* \Pi_i)^{1-\sigma} = Y_i/Y. \tag{18}$$

The left hand side of (18) is recognized as a behavioral share equation for the globally common CES preferences when all countries face a common world price $p_i^* \Pi_i$. This follows from dividing through (16) by E_j to give the representative CES demand share and then understanding that the CES price index that usually appears in the denominator is equal to one in the case of (18) because summing (18) implies $\sum_i (\beta_i p_i^* \Pi_i)^{1-\sigma} = 1$.

To complete the derivation of the structural gravity model, use (18) to substitute for $\beta_i p_i^*$ in (16), (17) and the CES price index. Then:

$$X_{ij} = \frac{E_j Y_i}{Y} \left(\frac{t_{ij}}{P_j \Pi_i} \right)^{1-\sigma} \tag{19}$$

$$(\Pi_i)^{1-\sigma} = \sum_j \left(\frac{t_{ij}}{P_j} \right)^{1-\sigma} \frac{E_j}{Y} \tag{20}$$

$$(P_j)^{1-\sigma} = \sum_i \left(\frac{t_{ij}}{\Pi_i} \right)^{1-\sigma} \frac{Y_i}{Y}. \tag{21}$$

A key component of the structural gravity model is multilateral resistance, the indexes P_j and Π_i that aggregate *all* bilateral trade costs, policy and non-policy, cross-border and internal. Multilateral resistance is interpreted as the demand (P_j) and supply (Π_i) side incidence of trade costs. Outward multilateral resistance Π_i aggregates the set of bilateral trade costs as if in effect country i shipped its product to a single world market at markup Π_i. This is the intuitive meaning of (18). For any bilateral trade cost, taking out the supply side incidence means that t_{ij}/Π_i is the buyer's bilateral incidence of the trade cost. Then P_j, inward multilateral resistance, is the the CES price index of the bilateral buyers incidences. It is as if the buyer in j goes to the world market and buys a bundle of goods, one from each seller, taking them home at the uniform markup P_j.

Since the system of equations (20)-(21) solves for $\{\Pi_i, P_j\}$ only up to a scalar,[10] an additional restriction from a normalization is needed. Relative multilateral resistances are what matters for resource allocation, so the normalization can be chosen for convenience in computation or interpretation — for example, $P_j = 1$ for some convenient reference country j.[11] See Anderson and Yotov (2008) for more details and an example of these methods applied to Canadian provincial trade.

The structural gravity model yields many other useful and intuitive indexes that decompose the incidence of trade costs. One is the uniform border policy index.

3.2 Uniform Border Policy

The uniform *policy* border barrier is defined by replacing the policy-related costs on all trade that crosses borders (i.e., excluding internal trade) with a border barrier for each country that is uniform across its trading partners subject to a constant value of the domestic value of international trade. The concept of uniform border barriers applied to *all* trade costs that act on borders, policy-related or not, was developed by Anderson and van Wincoop (2004) and discussed further in Anderson and Neary (2005, ch. 10).

The uniform border policy maintains the initial equilibrium demand and hence the initial world prices. It resembles the MTRI in being focused on international trade but it uses domestic as opposed to world prices to form the volume constraint. Like the compensated MTRI, it does not account for the effect on expenditure of changes in tariff revenue in going to the

[10]Notice that if $\{P_j^0, \Pi_i^0\}$ is a solution to (20)-(21) then so is $\{\lambda P_j^0, \Pi_i^0/\lambda\}$ for any $\lambda > 0$.

[11]With information on the factory gate prices p_i^* and the distribution parameters β_i, the natural normalization is $\sum_i(\beta_i p_i^* \Pi_i)^{1-\sigma} = 1$. The natural normalization is required when applying the model in a larger general equilibrium computational setting which includes reallocations across product classes. For reallocations across trading partners within a product class, all normalizations are equivalent.

hypothetical uniform barrier. In the context of evaluating total trade costs, of which trade policy is a small part, abstraction from redistribution of tariff revenue and quota rent is a perhaps justifiable simplifying assumption.

The switch to a uniform border policy potentially implies that a new set of multilateral resistances are generated by (20)-(21). However, the preservation of domestic value of trade along with the equilibrium market clearance implies that the value of internal shipments is constant. This in turn implies that multilateral resistances (and factory gate prices for producers) do not change. Constant multilateral resistances greatly simplify the calculation of the uniform border barrier.

With constant P's and Π's, the requirement that the hypothetical uniform import policies yield equal domestic value of international trade implies that

$$\sum_{i \neq j}(1+\tau_j^M)^{1-\sigma}(\bar{t}_{ij}/\Pi_i)^{1-\sigma}Y_i/Y = \sum_{i \neq j}(t_{ij}/\Pi_i)^{1-\sigma}Y_i/Y\,, \forall j \neq i, \tag{22}$$

where τ_j^M is the uniform border import policy (as an ad valorem tariff equivalent) and $\bar{t}_{ij}$ is the non-policy barrier that acts on the border. Solving (22) for $1+\tau_j^M$:

$$1+\tau_j^M = \left(\sum_{i \neq j}(t_{ij}/\bar{t}_{ij})^{1-\sigma}v_i^M\right)^{1/(1-\sigma)}, \forall j \neq i;$$

where

$$v_i^M = \frac{(\bar{t}_{ij}/\Pi_i)^{1-\sigma}Y_i/Y}{\sum_{i \neq j}(\bar{t}_{ij}/\Pi_i)^{1-\sigma}Y_i/Y}.$$

The uniform border policy formula has an intuitive structure. It is a CES index of the bilateral policy border barriers $t_{ij}/\bar{t}_{ij}$. The weights v_i^M in the index are based on the demand side bilateral incidences on non-policy barriers $\bar{t}_{ij}/\Pi_i$, normalized by their sum.

Symmetric to the uniform import policy formula is the uniform export policy formula. For each exporting country, define the uniform-across-countries border policy that maintains the domestic value of international trade. Then paralleling (22), the uniform export policy is implicit in:

$$\sum_{j \neq i}(1+\tau_i^X)^{1-\sigma}(\bar{t}_{ij}/P_j)^{1-\sigma}E_j/Y = \sum_{j \neq i}(t_{ij}/P_j)^{1-\sigma}E_j/Y\,, \forall i \neq j. \tag{23}$$

(23) yields a closed form solution for $1+\tau_i^X$:

$$1+\tau_i^X = \left(\sum_{j\neq i}(t_{ij}/\bar{t}_{ij})^{1-\sigma}v_j^X\right)^{1/(1-\sigma)}, \forall i \neq j;$$

where

$$v_j^X = \frac{(\bar{t}_{ij}/P_j)^{1-\sigma}E_j/Y}{\sum_{j\neq i}(\bar{t}_{ij}/P_j)^{1-\sigma}E_j/Y}.$$

τ_i^X is a concept that resembles the Kee *et al.* MA-OTRI, but it controls for potentially endogenous world prices in a single goods class.

It is straightforward to rank the uniform import border policy relative to the trade weighted average tariff: *the uniform border policy barrier exceeds the trade weighted average policy trade cost.* For proof, see Anderson and Neary (2005, ch. 10). The reason is the substitution effect — the trade weighted average under-weights high tariffs and over-weights low tariffs.

The ranking proposition implies that the amount by which the uniform border barrier exceeds the trade weighted average rises with the dispersion of the discriminatory policy. Thus for example EU members that individually have trade weighted average tariffs similar to the US are likely to have bigger uniform border barriers because they discriminate on a larger portion of their cross border trade than does the US.

3.3 Aggregating Across Sectors

It is natural to think of combining aggregation across trading partners with aggregation across sectors. The methods of the preceding section can be combined with the methods of this section. A survey is not the proper place for such a development, but the basic ingredients are here for an ambitious reader to combine. The essential trick is that the uniform border policy variable aggregates in such a way that world prices remain constant. Then it is possible to use the techniques of Section 2 to aggregate across sectors to form volume equivalent uniform tariffs. The compensated MTRI formed in this way is fully consistent. The uncompensated MTRI and the TRI involve complications that are beyond the scope of the survey; interested readers should consult Anderson and Neary (2005) for more information.

4 Conclusion

Measurement of protection is an important problem.Progress toward better solutions has been made but more is possible on both empirical and theoretical fronts. As for empirics, the opaque and often arbitrary nature of many non-tariff barriers points to inferential methods, buttressed by opinion surveys of businessmen who experience the barriers first hand. As for theory, better models lead to more reliable inference and to more appropriate and believable aggregation. The existing evidence points to large remaining trade barriers of the kind that policy can reduce, despite the fact that many formal tariffs are quite low.

5 References

Anderson, James E. and J. Peter Neary (1994), "The Trade Restrictiveness of the Multi-Fibre Arrangement" *World Bank Economic Review*, May 1994, 171-89.

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

Anderson, James E. and Eric van Wincoop (2004), "Trade Costs", *Journal of Economic Literature*, 42, 691-751.

Anderson, James E. and Yoto V. Yotov (2008), "The Changing Incidence of Geography", NBER Working Paper No. 14423.

Anderson, James E. (2009), "Consistent Trade Policy Aggregation", *International Economic Review*, 50(3), 903-27.

Anderson, James E. (1985), "The Relative Inefficiency of Quotas: the Cheese Case", *American Economic Review*, 75, 178-90.

Baier, Scott and Jeffrey Bergstrand (2007), Do free trade agreements actually increase members international trade?, *Journal of International Economics*, 71, 72-95.

Besedes, Tibor and Thomas J. Prusa (2003), "On the Duration of Trade", NBER Working Paper No. 9936.

Crozet Matthieu, Pamina Koenig and Vincent Rebeyrol (2008), CEPII Working Paper No. 2008 - 13.

Dixit, Avinash K. (2004), *Lawlessness and Economics*, Princeton: Princeton University Press.

Eaton, Jonathan and Samual Kortum (2002) "Technology, Geography and Trade," *Econometrica*, 2002, 70(5), 1741-1779.

Helpman, Elhanan, Marc J. Melitz and Shona Rubinstein (2008), "Trading Volumes and Trading Patterns", *Quarterly Journal of Economics*, forthcoming.

Irwin, Douglas (2007), "Trade Restrictiveness and Deadweight Losses from U.S. Tariffs, 1869-1961", NBER Working Paper No. 13450.

Kee, Hiau Looi, Alessandro Nicita and Marcelo Olarreaga (2008), "Import Demand Elasticities and Trade Distortions", *Review of Economics and Statistics*, 90, 666-682.

Kee, Hiau Looi, Alessandro Nicita and Marcelo Olarreaga (2009), "Estimating Trade Restrictiveness Indices", *Economic Journal*, 119, 172-99.

Trefler, Daniel (1993), "Trade Liberalization and the Theory of Endogenous Protection" *Journal of Political Economy*, 101, 138-60.