The International Agricultural Trade Research Consortium is an informal association of University and Government economists interested in agricultural trade. Its purpose is to foster interaction, improve research capacity and to focus on relevant trade policy issues. It is financed by United States Department of Agriculture (ERS, and FAS), Agriculture and Agri-Food Canada and the participating institutions.

The IATRC Working Paper series provides members an opportunity to circulate their work at the advanced draft stage through limited distribution within the research and analysis community. The IATRC takes no political positions or responsibility for the accuracy of the data or validity of the conclusions presented by working paper authors. Further, policy recommendations and opinions expressed by the authors do not necessarily reflect those of the IATRC or its funding agencies. For a complete list of IATRC Working Papers, books, and other publications, see the IATRC Web Site http://www.umn.edu/iatrc

A copy of this paper can be viewed/printed from the IATRC Web Site indicated above

*Dr. David W. Skully is an Economist with the U. S. Department of Agriculture, Economic Research Service.

Correspondence regarding this paper should be addressed to:

Dr. David W. Skully
USDA/ERS/MTED
Room 5168
1800 M Street NW
Washington, D.C. 20036-5831

May 1999
ISSN 1098-9218
Working Paper 99-6
1 Introduction

The objective of this paper is to examine the economics of tariff quota [TRQ] administration. Of the many methods of tariff quota administration, it attempts to distinguish between those most likely to bias and inhibit trade and those least likely to do so.

Section 2 provides an economic and legal introduction to tariff-rate quotas. It explains how TRQs operate and the rules governing TRQ administration in the GATT/WTO. In particular, it argues that GATT Article XIII evaluates TRQ administration by two criteria: non-discrimination and quota fill. These criteria of ‘good’ TRQ administration are employed throughout the paper.

Section 3 introduces the various forms of TRQ administration as defined by the WTO and shows how countries differ in their choice of administrative methods. It also presents the distribution of quota fill rates by method as reported to the WTO for 1995 and 1996, and explains why quota fill is not a good indicator of administrative performance.

Section 4 is the analytical core of the paper. It applies the two WTO criteria of ‘good’ TRQ administration to the various administrative methods. It opens with a summary of the conclusions of the section: some forms of administration are inherently superior. In fact, they can be ranked by their inherent risk of biasing trade (discrimination) and of underfill. If the reader has no compelling interest in the economic reasoning employed to rank the methods, one can skip the balance of section 4 and go directly to the conclusion. The balance of Section 4 first considers the superior efficiency of market allocation; that is, rationing a fixed supply among competing demands by market determined prices, be they auction bids or market clearing prices, with or without tariffs. Market allocation provides the basis for the evaluation of other allocation methods. The other methods are divided into two groups: quasi-market methods and discretionary methods. Quasi-market methods can be reduced to algorithms. This group includes first-come, first-served; license on demand; and historical allocation. A short case study of the U.S. Sugar TRQ is provided as an example of historical allocation. Discretionary methods cannot be reduced to algorithms: the discretion over the volume and sourcing of in-quota imports is delegated to one of two kinds of organization within the importing country, state trading organizations or producer organizations. A short case study of the Japanese Food Agency’s administration of the TRQ for wheat is provided as an example of import administration by a state trading organization.

Section 5 concludes the study.
2 TRQ Defined

2.1 An Economic Definition

A tariff quota is a two-tiered tariff. In a given period, a lower, in-quota tariff \( t \) is applied to the first \( Q \) units of imports and a higher over-quota tariff \( T \) is applied to all subsequent imports. The terms tariff quota and tariff rate quota are employed interchangeably in the literature; this paper follows this tradition. Technically, tariff quota is a more accurate description than tariff rate quota: narrowly interpreted, the former includes specific tariffs while the latter excludes them.

From a legal point of view, tariff quotas are not quantitative restrictions because they do not limit the quantity that may be imported. One may always import by paying the over-quota tariff. This opportunity is not available under a regular quota. However, if the over-quota tariff is prohibitive under normal market conditions it will yield exactly the same import volume as a traditional quota. The only condition that would cause a tariff quota to result in a different volume of trade than a standard quota is if the difference between the domestic price and the international price exceeds the over-quota tariff. At such a price difference it is profitable to import even after paying the over-quota tariff. Were a standard quota in place it would be legally impossible to expand the volume of imports beyond the quantitative restriction. Because of this (often slight) difference, a tariff quota is (in theory) less restrictive than a standard quota.

Figure 1 illustrates how a tariff quota influences the incentive to import.1 The effective supply curve of exports to the import market consists of two horizontal lines. The lower line represents the in-quota imports and extends from 0 to \( Q \) at the price \( 1+t \). The other line represents the effective supply of over-quota imports and extends from \( Q \) to infinity at the price \( 1+T \). At import volume \( Q \) there is a discontinuity: a vertical line joins the in-quota and over-quota segments.

The effect of a tariff quota on trade depends on the excess demand for imports. The figure shows four possible excess demand conditions. Excess demand curves labeled 1 through 4 represent increasing levels of import demand.

1 no trade occurs: \( M_1 = 0 \).
2 quota is not binding: \( M_2 < Q \).
3 quota is binding: \( M_3 = Q \).
4 quota no longer binding: \( M_4 > Q \).

In the first case domestic excess demand is not sufficient to support imports at the world price, even without the in-quota tariff, so imports are zero: \( M_1 = 0 \). In the second case excess demand is

---

1 Throughout this paper we assume that the international market is competitive and that the importing country employing the tariff quota is a ‘small country’ in the sense that changes in its import volume are not sufficient to change international prices. These assumptions are reasonable for most existing tariff quotas and they allow us to represent the international supply curve as a horizontal line. To simplify matters further we express all prices in terms of the relevant international free market price \( W \). Consequently the international price is always equal to \( 1 \{W/W\} \) and the domestic price \( P = P/W \). This assumption also allows us to ignore the distinction between specific and ad valorem tariffs.
Figure 1  TRQ and Import Demand

Rent
sufficient to result in imports of M2 but not so great as to cause the quota to be binding: M2<Q. As long as imports satisfy domestic excess demand at a volume less than Q, the tariff quota functions as an ordinary tariff applied at the in-quota rate.

The tariff quota is binding in the third case. If there were no tariff quota and merely a tariff applied at the in-quota rate, imports of T3 would result (T for tariff). [Although they are not graphed T1 = M1 and T2 = M2]. Were the tariff applied at the rate of zero, imports of F3 would result (F for free trade). So one may write M3 = Q < T3 < F3. That imports when a TRQ is binding are less the volume of imports at the unconstrained in-quota tariff means that M3 units of imports must be rationed among T3 units of demand. Rationing is the essence of TRQ administration.

Markets ration available supplies among willing demanders through adjusting prices. This is textbook economics: the equilibrium price equates supply and demand. Tariff quotas can be administered using prices and market mechanisms. For example, one can cause imports to equal Q by increasing the in-quota tariff to P-1. There are at least three problems with this solution. First, it requires rather exact knowledge of domestic excess demand elasticity and other market information to determine the right tariff. Second, market conditions can change quickly, so even if one did get the tariff right, it would likely not stay right for long. Finally, even if one could always determine the ideal tariff, varying a tariff to keep it at the right level amounts to a form of variable levy and variable levies are prohibited by the WTO. (URAA, Annex 5.2)

The opportunity to import within the tariff quota and sell on the domestic market is an opportunity to gain a riskless profit on each item so imported. This rent is equal to the difference between the domestic price and the world price with the in-quota tariff: R = P - (1+t). Raising the tariff to R+t, as described in the previous paragraph, amounts to taxing away the rent. As noted above, this is hard to do. However, one can establish a market for the right to import within the quota. If the right to import within the quota were auctioned potential importers would be willing to pay approximately R for the opportunity to make a riskless profit of R.

Excess demand curve 4 represents a high level of demand, sufficient to sustain imports at the over-quota tariff. The volume of imports is no longer constrained at Q. The domestic price increases to P=1+T. However, the rationing problem remains for the volume within the quota. The opportunity to import at 1+t and sell for the domestic price of 1+T is still available for Q units of imports. Rent equals T-t = (1+T)-(1+t), the maximum possible under a TRQ. The rent still must be rationed.

2.2 TRQ Administration and the WTO: A legal definition

Tariff quota administration concerns how the rights to import at the in-quota tariff are distributed. How these rights are distributed can determine both the volume and distribution of trade as well as the distribution of quota rents. When considering tariff quotas it is important to keep clear the

---

2 The shaded triangle represents the dead weight welfare loss of the tariff quota relative to an applied tariff at the in quota rate. An extensive welfare analysis is pursued in section 2.
distinction between the volume and distribution of trade and the volume and distribution of rents. The WTO is only concerned with how quota administration influences the volume and distribution of trade; it has no direct interest in the distribution of rents. However, it is the distribution of rents that motivates the politics of TRQ administration. The choice of the method of tariff quota administration is a political decision; many competing interests claim entitlement to quota rents.

Tariff quota administration is fundamentally a rationing problem. There are many ways to ration something the issue is to determine whether some ways are better than others.

**GATT Article XIII and Distributive Justice**

GATT Article XIII governs the administration of quantitative restrictions [QRs]. QRs were effectively prohibited in the Uruguay Round. Article XIII survives to govern TRQs which, while legally and technically not QRs, generally function as if they were QRs. Moreover TRQs pose administrative problems identical to QRs. The next few paragraphs provide a brief history of the treatment of TRQs in international trade agreements. In particular, the narrative focuses on underlying conflict between the GATT principle of non-discrimination and the rationing problem posed by TRQs.

Quantitative restrictions, such as absolute quotas and tariff quotas, were rarely applied to imports until after World War I. The League of Nations, in the series of World Economic Conferences it sponsored in the inter-war period, attempted to reconcile how QRs could be administered without discrimination, that is, consistent with Most Favored Nation [MFN] principles. By 1930 four positions could be discerned in these deliberations:

1. QRs are *per se* inconsistent with MFN.
2. MFN requires that each country be assigned an equal share of the global quota.
3. MFN can be approximated by allotting the global quota in proportion to the trade shares of current suppliers.
4. Allow the global quota to be filled on a first come first served basis.

Because of the conflicting interpretations of the principle of non-discrimination, there was no consensus, except that “there was fairly unanimous agreement that the use of global, race-to-the-border quotas (now permitted by GATT Article XIII) was inconsistent with MFN because it unduly favored countries with geographical proximity and/or better transport facilities.” [Hudec (1997): 178, n. 14]

Young (1994) argues that any rationing system consists of three kinds of decision: a supply decision; a distributive decision; and a reactive decision. Together this sequence of three decisions results in an effective decision or allocation. In the case of quotas or tariff quotas, the supply decision is made by the importing country in its tariff schedule. The distributive decision is a choice among rules for allocating the limited supply among competing claimants. The supply and distributive decisions are made by institutions; reactive decisions are made by individual claimants. Claimants attempt to
optimize given the incentives and the constraints on action created by the supply and distributive
decisions. TRQ administration belongs to the class of distributive decisions or allocation rules.

The allocation rules that we see in practice usually exhibit one of three broad conceptions of
equity. Parity means that the claimants are treated equally, either because they actually are
equal or because there is no clear way to distinguish among them. Proportionality
acknowledges differences among the claimants and divides the good in proportion to these
differences. Priority asserts that the person with the greatest claim to the good gets it. [Young
(1994: 8)]

These three general principles of distributive justice are evident in the four positions on non-
discriminatory QR allocation above. The first position claims there is no just way to solve the quota
allocation problem. The second position argues for strict parity: if there are N parties to a trade
agreement then a fair allocation gives each party exactly 1/N of the global quota. The third position
advocates proportionality, and that the just basis for proportionality is the observed volume of trade
in some recent representative period. Finally, the fourth position asserts (literal) priority in the form
of first come, first served. As the brief history below shows, neither the League of Nations, nor the
Havana Charter, nor the GATT resolved this issue. Instead of advocating one principle of distributive
justice and proscribing all others, Article XIII allows a conflicting set of distributive principles to co-
exist. Predictably, this leads to trade conflicts over TRQ administration.

Starting in the early 1920s, political demands for agricultural protection and intervention emerged in
Europe and North America. Britain and the Netherlands remained resolute free traders until the
passage of the U.S. Smoot-Hawley Tariff Act in 1930. After that, protectionism cascaded: In 1932
Britain concluded the Ottawa Agreement that established a system of ‘imperial preference.’ Britain
increased its MFN tariffs but granted a margin of preference to imports from its imperial dominions;
the dominions, in turn, also increased MFN tariffs but granted reciprocal preferences to Britain. The
Smoot-Hawley Act raised tariffs on imports from all sources. By 1934, with the passage of the
Reciprocal Trade Agreements Act [RTAA], the United States had shifted to a trade strategy of
‘discriminatory liberalization.’ The RTAA generated a network of bilateral preferences between the
United States and selected trading partners. By the outbreak of the Second World War, Belgium,
Britain, France, Germany, Italy, Japan, the Netherlands, and the United States, among others, had
developed systems of discriminatory trade preferences. The allied powers attempted to maintain
these systems after the war by incorporating them into the Post-War order. The GATT, one of the
three pillars of the Post-war International Economic order, required devices such as Article XIII to
preserve these obvious violations of the principle of non-discrimination.4

When governments decided, after World War II, that QR’s would be permitted in many

3 See Skully (1993) on competing visions of the governance of agricultural trade in the 1940s.
4 The three pillars were to be: IMF, IBRD and ITO. The ITO was aborted when the Havana Charter was not ratified
by the United States. The GATT is the remnant of the ITO. Article 22 of the Havana Charter addresses the
administration of QRS, most of the text was incorporated into GATT Article XIII.
situations. ... It became necessary to say, whether it was true or not, that QR’s could be applied in a manner consistent with the MFN concept. And so GATT Article XIII was written. Given its less-than-robust conceptual heritage, it is a small wonder that Article XIII proved to be a rather sickly child. [Hudec (1997: 178)]

Article XIII is a ‘sickly child’ because of a congenital deformity: it advocates both non-discrimination and discrimination. The interpretation of Article XIII which follows emphasizes its advocacy of the principle of non-discrimination, the principle of distributive justice upon which the GATT is constructed. Later, in Sections 2 and 3, the sub-paragraphs of Article XIII that allow for discrimination are considered in depth.

**Article XIII and Non-Discrimination**

Article XIII of the GATT, “Non-Discriminatory Administration of Quantitative Restrictions,” governs the administration of tariff quotas. A reading of Article XIII and other, related, documents leads one to conclude that the GATT advocates two criteria for judging whether tariff quotas are being properly administered: 1) quota fill and 2) distribution of trade.

*Quota fill* requires that imports of the in-quota volume be allowed if market conditions permit. That is, tariff quota administrators should not impose any impediments to imports beyond payment of the in-quota tariff. Quota fill implies a two-part test. One first asks: Was the quota filled? This is a relatively simple empirical issue. If the quota is not filled, the second question is posed: Did market conditions permit imports? Figure 1 also illustrates what ‘market conditions permit’ means in this context. Suppose demand curve 3 represents normal import demand, but there is an unusually large domestic harvest. Domestic supply expands and reduces excess demand for the product. The new excess or import demand, 1, is drawn so that the domestic market clearing price equals 1, the world price, and imports are zero. The quota is not filled, but there is a legitimate reason: there is simply not sufficient domestic demand under current market conditions. Were domestic demand to increase slightly to 2, imports would satisfy excess demand, partially filling the quota. This too, is an instance of legitimate underfill. In practice, the simplest second-stage test is to ask whether the domestic price is less than the world (border) price plus the in-quota tariff. If so, there is clearly no excess demand for imports. If the domestic price is greater than 1+t and the quota is not filled, it indicates that profitable arbitrage opportunities are not being realized. It is then appropriate to inquire why. There may be legitimate transactions costs that the analyst has overlooked. But it may be because of the method of TRQ administration.

*Distribution of trade.* GATT Article XIII paragraph 2 states:

2. In applying import restrictions to any product, contracting parties shall aim at a distribution of trade in such product approaching as closely as possible the shares which the various contracting parties might be expected to obtain in the absence of such restrictions…
This language implies the construction of a free trade or tariff equivalent counterfactual. That is, one determines what the distribution of trade (supplier market shares) would be were a tariff employed to restrict imports to equal \( Q \). The administration of the tariff quota is then evaluated by how closely the distribution of the restricted volume of trade (under tariff quota) approaches the counterfactual distribution. Basically, were one to construct a pie chart of supplier market shares under a nondiscriminatory tariff quota, it should look identical to the tariff equivalent pie chart. The economic principle underlying the distribution of trade criterion is the minimization of trade distortions given the tariff quota constraint. The GATT principal of nondiscrimination asserts that trade shares should be determined by the relative efficiency of suppliers and not by alternative, discriminatory criteria. Clearly, *market allocation* is the principal of distributive justice governing TRQ administration under Article XIII:2: what this means and how it may be operationalized is discussed in Section 2.2.

The two criteria result in two evaluative questions: 1) Which methods of allocation are most likely to result in quota underfill? And 2) Which methods of allocation are most likely to result in a discriminatory distribution of trade?

### 3 Choice and Performance of TRQ Administration

#### 3.1 Patterns of Administrative Choice

The World Trade Organization identifies seven principal methods of TRQ administration. Member nations are to notify WTO whether and how the tariff quotas listed in their tariff schedules are being administered. Of the 1278 notified tariff quotas in 1996, over half were not enforced; rather all imports were allowed at the in-quota tariff. However, the over-quota tariff can be re-applied at will. Of the 632 TRQs that were enforced, almost half were allocated by license on demand. The box lists the seven methods ordered by frequency of use (in percent).

Among the countries that have notified TRQs to the WTO there are some clear patterns of administrative preference. Tables 1 and 2 display country preferences: table 1 shows the percentage distribution of administrative modes, while table 2 shows the absolute number of tariff quotas. A concentration index\(^5\) is also displayed: 1.00 indicates that all TRQs share the same method of administration, the lower the index the greater the variety of methods employed.

### Applied Tariffs

Four countries treat all notified TRQs as applied tariffs and ten more treat over half as applied tariffs. Regional trade agreements partially explain membership in this set of countries. Norway and Iceland

---

\(^5\) The Index is the Hirschman-Herfindahl measure: \( H = \sum w_i^2 \). Where \( w_i \) is the percentage share of a country’s TRQs administered by method \( i \).
Table 1

Proportion of TRQs notified by country, by method

<table>
<thead>
<tr>
<th>Administrative Method</th>
<th>AT</th>
<th>LD</th>
<th>HI</th>
<th>AU</th>
<th>FC</th>
<th>ST</th>
<th>PG</th>
<th>OT</th>
<th>MX</th>
<th>NS</th>
<th>Index</th>
<th>WTO code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Czech Rep</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CZE</td>
</tr>
<tr>
<td>Morocco</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MAR</td>
</tr>
<tr>
<td>US</td>
<td>0.68</td>
<td>0.07</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>USA</td>
</tr>
<tr>
<td>EU</td>
<td>0.65</td>
<td>0.20</td>
<td>0.10</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EEC</td>
</tr>
<tr>
<td>Japan</td>
<td>0.50</td>
<td>0.17</td>
<td>0.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>JPN</td>
</tr>
<tr>
<td>Indonesia</td>
<td>0.09</td>
<td>0.48</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IND</td>
</tr>
<tr>
<td>Thailand</td>
<td>0.17</td>
<td>0.33</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ISR</td>
</tr>
<tr>
<td>Korea, Rep</td>
<td>0.31</td>
<td>0.25</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>KOR</td>
</tr>
<tr>
<td>Canada</td>
<td>0.05</td>
<td>0.29</td>
<td>0.24</td>
<td>0.19</td>
<td>0.05</td>
<td>0.05</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CAN</td>
</tr>
<tr>
<td>Switzerland</td>
<td>0.18</td>
<td>0.14</td>
<td>0.07</td>
<td>0.14</td>
<td>0.21</td>
<td>0.00</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CHE</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CRI</td>
</tr>
<tr>
<td>Barbados</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BRB</td>
</tr>
<tr>
<td>New Zealand</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NZL</td>
</tr>
<tr>
<td>El Salvador</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SLV</td>
</tr>
<tr>
<td>Ecuador</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ECU</td>
</tr>
<tr>
<td>Venezuela</td>
<td>0.97</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>VEN</td>
</tr>
<tr>
<td>Norway</td>
<td>0.92</td>
<td>0.03</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NOR</td>
</tr>
<tr>
<td>Mexico</td>
<td>0.91</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MEX</td>
</tr>
<tr>
<td>Nicaragua</td>
<td>0.89</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NIC</td>
</tr>
<tr>
<td>Iceland</td>
<td>0.87</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ISL</td>
</tr>
<tr>
<td>Poland</td>
<td>0.73</td>
<td>0.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>POL</td>
</tr>
<tr>
<td>Guatemala</td>
<td>0.70</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GTM</td>
</tr>
<tr>
<td>Romania</td>
<td>0.58</td>
<td>0.33</td>
<td></td>
<td></td>
<td></td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ROM</td>
</tr>
<tr>
<td>R.South Africa</td>
<td>0.60</td>
<td>0.38</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ZAF</td>
</tr>
<tr>
<td>Colombia</td>
<td>0.49</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td>COL</td>
</tr>
<tr>
<td>Malaysia</td>
<td>0.05</td>
<td>0.89</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MYS</td>
</tr>
<tr>
<td>Hungary</td>
<td>0.03</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HUN</td>
</tr>
<tr>
<td>Slovakia</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SVK</td>
</tr>
<tr>
<td>Slovenia</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SVN</td>
</tr>
<tr>
<td>Australia</td>
<td>0.50</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AUS</td>
</tr>
<tr>
<td>Philippines</td>
<td>0.36</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PHL</td>
</tr>
<tr>
<td>Tunisia</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TUN</td>
</tr>
</tbody>
</table>
### Table 2

**Number of TRQs notified by country, by method**

<table>
<thead>
<tr>
<th>Country</th>
<th>TRQs</th>
<th>Tariff</th>
<th>License</th>
<th>Historical</th>
<th>Auction</th>
<th>First come</th>
<th>First serve</th>
<th>State</th>
<th>Producer</th>
<th>Other</th>
<th>Mixed</th>
<th>Not Specified</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech Rep</td>
<td>24</td>
<td>24</td>
<td>1</td>
<td>17</td>
<td>5</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Morocco</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>US</td>
<td>32</td>
<td>20</td>
<td>6</td>
<td>17</td>
<td>5</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.46</td>
</tr>
<tr>
<td>EU</td>
<td>58</td>
<td>6</td>
<td>20</td>
<td>17</td>
<td>5</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.53</td>
</tr>
<tr>
<td>Japan</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.48</td>
</tr>
<tr>
<td>Indonesia</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.50</td>
</tr>
<tr>
<td>Thailand</td>
<td>2</td>
<td>11</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.32</td>
</tr>
<tr>
<td>Israel</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.24</td>
</tr>
<tr>
<td>Korea, Rep</td>
<td>21</td>
<td>17</td>
<td>6</td>
<td>10</td>
<td>3</td>
<td>10</td>
<td>10</td>
<td>22</td>
<td>30</td>
<td>0.31</td>
<td>0.61</td>
<td></td>
<td>0.22</td>
</tr>
<tr>
<td>Canada</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.19</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.55</td>
</tr>
<tr>
<td>Barbados</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>New Zealand</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>El Salvador</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Ecuador</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Venezuela</td>
<td>59</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.94</td>
</tr>
<tr>
<td>Norway</td>
<td>213</td>
<td>1</td>
<td>8</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td>10</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nicaragua</td>
<td>8</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iceland</td>
<td>78</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poland</td>
<td>79</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guatemala</td>
<td>16</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Romania</td>
<td>7</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R. South Africa</td>
<td>32</td>
<td>20</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>0.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td>33</td>
<td>33</td>
<td>1</td>
<td>33</td>
<td>1</td>
<td></td>
<td></td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malaysia</td>
<td>1</td>
<td>17</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>2</td>
<td>68</td>
<td>License on</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slovakia</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slovenia</td>
<td></td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Philippines</td>
<td>5</td>
<td>9</td>
<td>Historical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tunisia</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Index**

Europe Candidates: CZE CHE NOR ISL POL ROM HUN SVK SVN

CACM Members: CRI BRB SLV ECU VEN MEX NIC GTM COL
are both members of the European Economic Area. The EEA is not a customs union; each member maintains a distinct tariff schedule. EEA members attempt to harmonize policies with the EU, but agricultural products are excepted. Iceland and Norway together account for 291 TRQs, almost one-fourth of the global total. However, only 31 are enforced through non-tariff means. Auctions are used to allocate minimum access tariff quotas, and Norway makes some historical allocations. The large number of tariff quotas administered as applied tariffs allows these two countries to enjoy a relatively liberal trade regime (given the tariff levels applied) while preserving the option to enforce the tariff quotas in a more protective manner. The liberty to revert to a more protective trade regime would likely allow them to join the EU and harmonize their agricultural policies without having to compensate third countries.

<table>
<thead>
<tr>
<th>Method</th>
<th>Brief Explanation</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Applied Tariff</td>
<td>Unlimited imports are allowed at or below the in-quota tariff rate.</td>
<td>646</td>
<td>50.5%</td>
</tr>
<tr>
<td>2 License on Demand</td>
<td>Licenses are required to import at the in-quota tariff. If the demand for licenses is less than Q, this system operates as a First-Come, First-Served system. If demand is greater than Q, the import volume requested is reduced proportionately among all applicants.</td>
<td>314</td>
<td>24.6%</td>
</tr>
<tr>
<td>3 First-Come, First-Served (FCFS)</td>
<td>The first Q units of imports to clear customs are charged the in-quota tariff, all subsequent imports are charged the over-quota tariff.</td>
<td>104</td>
<td>8.1%</td>
</tr>
<tr>
<td>4 Auction</td>
<td>The right to import at the in-quota tariff is auctioned.</td>
<td>76</td>
<td>5.9%</td>
</tr>
<tr>
<td>5 Historical</td>
<td>The right to import at the in-quota tariff is allocated to firms on the basis of their trading volume in previous periods.</td>
<td>30</td>
<td>2.3%</td>
</tr>
<tr>
<td>6 State Trading</td>
<td>The right to import at the in-quota tariff is granted wholly or primarily to a state trading organization.</td>
<td>22</td>
<td>1.7%</td>
</tr>
<tr>
<td>7 Producer Group</td>
<td>The right to import at the in-quota tariff is granted to an organization representing producers, processors or distributors of the controlled product.</td>
<td>8</td>
<td>0.6%</td>
</tr>
<tr>
<td>Mixed</td>
<td>A combination of two or more of the seven methods above.</td>
<td>47</td>
<td>3.7%</td>
</tr>
<tr>
<td>Other</td>
<td>Methods which do not correspond to the seven methods above.</td>
<td>21</td>
<td>1.6%</td>
</tr>
<tr>
<td>Not Specified</td>
<td>The method is not specified in the WTO notification</td>
<td>10</td>
<td>0.8%</td>
</tr>
</tbody>
</table>

Costa Rica, El Salvador, Nicaragua and Guatemala are members of the Central American Common Market [CACM]. The CACM adopted a common external tariff [CET] in 1993. Nicaragua and
Honduras are allowed a longer transition for harmonizing with the CET; this may account for the absence of Honduran TRQs. Similarly, within the Andean Community, Colombia, Ecuador and Venezuela formed a Customs Union with a CET in 1995. Bolivia and Peru are also Andean members but retain separate tariff schedules. Finally, Mexico has established bilateral agreements with several of Andean and CACM members.

**License on Demand**

Romania and the five member nations of the Central European Free Trade Association [CEFTA] -- Poland, Slovakia, Czech Republic, Hungary, and Slovenia -- have also adopted relatively similar tariff quota administration methods. The Czech Republic is the outlier of this group having adopted FCFS for all TRQs. Slovakia, Slovenia and Hungary reveal a strong preference for License on Demand; and it is the second most common method for Poland and Romania. Perhaps this a legacy of the Habsburg Empire. As each of these countries has applied for EU accession, they, like Norway and Iceland, have an interest in maintaining the potential to enforce more protectionist border measures.

**Diverse Methods**

Nine countries can be distinguished by the wide variety of administrative methods they employ. With the exception of Indonesia and Thailand, this group includes the major industrialized countries with protectionist agricultural policies. The proportion of applied tariffs in this group is less than one-fifth, and several countries have no applied tariff TRQs. This is the only group to rely heavily on State Trading, Producer Groups, Mixed and other Methods, indicative of the complicated political economy of protection in industrial economies.

### 3.2 Patterns of Quota Fill

What can be inferred from the analysis of quota fill rates? (See Table 3) The WTO Secretariat routinely reports the fill rates of notified TRQs. The reports calculate the average quota fill rates by administrative method and commodity group (dairy, grain, etc.) among other factors. The reports do not go beyond reporting these basic descriptive statistics. They refrain from drawing inferences from quota fill data about the relative merits of the various administrative methods. Such restraint is appropriate. It is tempting to draw conclusions from quota fill statistics; but caution is advised because the data are fraught with hazards. For example, quota fill rates are significantly higher for historical allocation and state trading organizations than for auction allocation. This is the opposite of what this paper argues in section 4. The author would readily concede that this paper’s theoretical deductions were flawed if the quota fill data were the result of a properly designed test of the relative risks of underfill and distribution bias of the various administrative methods. A proper test would require, among other things, randomly assigning administrative methods among countries and commodities. However, there is a political economy of administrative choice: governments choose which administrative method to employ for each commodity. Thus, fill rates reflect factors determining a government’s choice of method as well as the intrinsic properties of the administrative method employed and the commodity market conditions during the period of observation. It is
Table 3

Distribution of Quota Fill Rates by Administrative Method

<table>
<thead>
<tr>
<th>Administrative Method</th>
<th>Distribution of TRQs by fill-rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Trader</td>
<td>![Graph showing distribution]</td>
</tr>
<tr>
<td>Historical</td>
<td>![Graph showing distribution]</td>
</tr>
<tr>
<td>Producer Group</td>
<td>![Graph showing distribution]</td>
</tr>
<tr>
<td>First Come, First Served</td>
<td>![Graph showing distribution]</td>
</tr>
<tr>
<td>License on Demand</td>
<td>![Graph showing distribution]</td>
</tr>
<tr>
<td>Auction</td>
<td>![Graph showing distribution]</td>
</tr>
<tr>
<td>Applied Tariff</td>
<td>![Graph showing distribution]</td>
</tr>
<tr>
<td>Avg Non-Applied Tariffs</td>
<td>![Graph showing distribution]</td>
</tr>
</tbody>
</table>

- □ below 20%
- □ 20% - 80%
- ■ Above 80%
difficult to identify and separate these factors. The following paragraphs attempt to sketch a political economy of administrative choice approach to the quota fill data.

Very simply, the more politically sensitive imports of a commodity are the greater the probability administration will be by discretionary methods. In the Uruguay Round many WTO members agreed to construct minimum access TRQs to allow a specified proportion of domestic consumption to be imported. For countries where imports had been banned previously, such minimum access TRQs present a serious challenge to domestic producers. In almost all cases domestic prices greatly exceeded world prices and it was obvious the TRQ would fill. Such TRQs have often been administered by either State Trading Organizations or Producer Groups. By delegating the import authority, in part or in whole, to the domestic industry the industry can minimize or manage the competitive threat of imported products. By this reasoning one should expect to find that discretionary methods are most likely to be applied to TRQs which are likely to fill. Moreover, discretionary methods tend to attract the scrutiny of potential exporters and their governments, so these TRQs are especially well policed and, as a consequence, are generally enforced to the letter of the law.

A corollary is that for imports that are not particularly sensitive there is less need to manage them and thus it is more likely that in-quota imports will be administered by market or quasi-market methods, or by the most liberal means, by not imposing the over-quota tariff at all, that is, by the ‘applied tariff’ method. Similarly, for TRQs that are unlikely to fill because the domestic price is generally below the world price, the risk of import disruption is low and so is the need to manage imports through discretionary means.

4 Economic Analysis of TRQ Administrative Methods

4.1 Summary of the Analysis

The ranking of methods is outlined in the box on the next page.

**Market methods**: The rationale behind the ranking follows from the mechanisms employed to allocate tariff quota import rights. If the basis for judging good tariff quota administration is the tariff equivalent counterfactual, then the tariff equivalent *Applied Tariff* must be the best method for allocation. Of course, an applied tariff is not a TRQ, so it is not a viable solution to the TRQ rationing problem. Auctioning tariff quota rights results in a volume and distribution of trade equal to the applied tariff equivalent to the quota. Consequently, *Auctions* have the least risk of bias. If the transactions costs of auctioning quotas are minimal, the risk of underfill is also minimal. The source of the efficiency of auctions and applied tariffs is their reliance on market determined prices to match least cost suppliers with those demanders with the greatest willingness to pay. The other allocation methods deviate from market based allocation.
**Quasi-Market methods:** The second tier of allocation methods consists of **License on Demand** and **First-Come, First-Served**. These methods differ from market based allocation because they add a random element to the market allocation process. Neither system discriminates between efficient and inefficient suppliers as effectively as an auction or a tariff. Because less efficient suppliers have a chance of gaining a share of the market they demand licenses or enter first-come, first-served queues. If they are lucky they gain a share of the quota and displace a lower cost supplier, thus biasing the distribution of trade from its tariff equivalent norm. The uncertainty and transactions costs introduced by these allocation methods can, at the margin, inhibit imports, thus increasing the probability of underfill.

**Historical Allocation** is an extreme version of a quasi-market method. It grants quota rights to importing firms or exporting countries (and sometimes both) on the basis of their shares of the trade in the controlled product in some period prior to the imposition of the tariff quota. Historical allocation is a superficially attractive method. First, by locking in the *status quo ante* distribution of trade it satisfies naive claims to quota entitlement. Second, the method results in numerical targets that are easy for officials to monitor and enforce. Because market conditions change, often rapidly, historical allocation is unlikely to result in a distribution of trade similar to the counterfactual norm; and the longer the shares remain fixed the greater the risk and magnitude of the bias. Moreover, in many cases, the *status quo ante* employed for historical allocation is not one generated by market forces. Consequently historical allocation perpetuates historical biases.

<table>
<thead>
<tr>
<th>Method</th>
<th>Risk of Under-fill</th>
<th>Risk of Biased Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Market Allocation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applied Tariff</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Auction</td>
<td>Low</td>
<td>Least</td>
</tr>
<tr>
<td><strong>Quasi-market Allocation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>License on Demand</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>First-Come, First-Served</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Historical</td>
<td>Moderate</td>
<td>Very High</td>
</tr>
<tr>
<td><strong>Discretionary Allocation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State Trading</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Producer Group</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

**Discretionary Methods:** The third tier of allocation methods is comprised of methods that delegate import sourcing decision power to special groups in the importing country. **State Trading Organizations** and **Producer Groups** are granted the decision making rights of what, when and from whom to import, given the constraints specified by the TRQ and market access commitments to the WTO. While it is theoretically possible that the managers of these organizations could make
decisions that replicate those generated by a market, in practice, they have other objectives and the pattern of trade will differ predictably from the counterfactual norm. As for the potential of underfill, if these organizations bias the volume of imports, they are generally likely to underfill unless carefully monitored by export suppliers.

Caveat. Please note that this ranking of methods only holds if trading in quota rights is prohibited. Allowing allocated quota rights to be traded in secondary markets can reduce many of the inefficiencies caused by the primary allocation. In practice, trading quota rights tends to be prohibited or narrowly proscribed by quota administrators. The issue of quota trading is discussed below.

4.2 Rationing and Markets

Tariff quota administration amounts to rationing. Given that one must ration, there are better and worse ways of doing it. This section examines in depth the assets and liabilities of the seven basic methods employed to administer tariff quotas. In particular, it evaluates them according to the WTO/GATT criteria of quota fill and distribution of trade.

Modern economics has largely ignored how markets actually operate. Granted, economists are always shifting supply and demand curves about, but they rarely focus on the logistics of how buyers and sellers actually find each other and how they arrive at prices so that supply equals demand. Adam Smith (1776) alluded to the universal human propensity to truck and barter. Leon Walras (1874) imagined an auctioneer who adjusts prices until markets clear. Curiously the empirical analysis of market micro-structure, how the institutions and rules governing exchange influence the efficiency of exchange, is a relatively recent phenomenon. Starting with the work of Vernon Smith (1962), economists have come to view the micro-structure of most markets as double auctions. Empirical studies of how trading rules and access to information influence market performance started with laboratory studies. But it has been financial economists who have truly embraced the topic, resulting in numerous studies of trading performance in stock and commodity exchanges, among other applications.6 Such studies are influencing the rules governing the NASDAQ and the design of electronic trading systems such as GLOBEX. In addition, they are causing economists to re-ask fundamental questions such as: “What Makes Markets Allocationally Efficient?” The quotation below comes from an article of just that title by Gode and Sunder (1997).

Allocative efficiency is the ratio of the actual to the potential gains from trade, which are equal to the sum of Marshallian consumer and producer surplus. Allocative efficiency is high if the consumers who value a good the most are able to buy it from the lowest cost producers. Consequently market designers and researchers want to know what determines the allocative efficiency of markets. Of particular interest to them is how market rules influence interactions among market participants and thereby affect efficiency. [603]

---

6 O’Hara (1995) is the standard survey; on NASDAQ see Schwert (1997) and citations therein; Carlton (1991) provides a Coasian perspective; Workshop on Double Auctions (1991) is the proceedings of a Sante Fe Institute conference on the topic.
A market generates maximum surplus if all buyers and sellers to the left of the intersection of supply and demand curves can find each other and exchange. Such buyers are inframarginal buyers: their willingness to pay exceeds the market clearing price. Extramarginal buyers, represented on the demand curve right of equilibrium, have a willingness to pay less than the market clearing price. Similarly, inframarginal suppliers have a willingness to accept less than the market clearing price; extramarginal suppliers, represented on the supply curve right of equilibrium, have a willingness to accept which exceeds the market clearing price.

Gode and Sunder identify three potential sources of inefficiency in exchange, two of which are relevant for the study of TRQs: that inframarginal buyers and sellers fail to find each other and negotiate a trade; and that extramarginal traders displace inframarginal traders. The former problem involves a failure to communicate or convey information. The latter involves the failure to discriminate between inframarginal and extramarginal trading partners.

The displacement of inframarginal traders by extramarginal traders is the primary source of inefficiency in TRQ administration. Indeed, inframarginal displacement is the fundamental theme of this paper. Market mechanisms minimize the risk of displacement. The further one deviates from market mechanisms, the greater the risk of displacement, the greater the risk of underfill, and the greater the risk of biased trade distribution.

**Welfare Analysis of the Rationing Problem**

Figure 2 reproduces the essential features of Figure 1, but includes more information for use in the comparative welfare analysis of alternative methods of TRQ administration. To review: the international supply curve is a horizontal line at 1. The in-quota tariff \( t > 0 \) is applied to the first \( Q \) units of imports. This shifts the effective import supply curve upwards to 1+t until the volume \( Q \) is attained. At the quota volume \( Q \) there is a vertical jump in the supply curve. Imports in excess of \( Q \) are charged the over-quota tariff \( T > t \) so, the effective supply curve continues at 1+T.

Figure 2 illustrates the case where import demand is sufficient to fill the quota, but not so great as to import at the over-quota tariff. [Demand curve 3 in figure 1] Consider the differences between a tariff quota, a simple tariff and free trade. With free trade an unlimited quantity can be imported at the world price. The domestic market clears with imports of \( Q \) and the domestic price equal to the world price, \( P = W = 1 \), and all demand inframarginal to \( P = 1 \) is satisfied. The large triangle below the demand curve and above the supply curve is the economic surplus gained from international exchange: that is, the sum of the areas labeled: \( A + Rent + Tariff + B + C + D \).

Under a simple tariff unlimited imports are allowed at the in-quota tariff \( t \). The domestic market clears with imports of \( Q \) and the domestic price is 1+t. Demand inframarginal to \( P = 1+t \) is satisfied. Domestic consumers= surplus equals the area: \( A + Rent + B \) and the domestic government collects the area \( Tariff + C \) in tariff revenue. Triangle \( D \) is the deadweight welfare loss from imposing the
Figure 2  TRQ Welfare Analysis

- Domestic Price / World Price
- $1 + T$
- Domestic Price
- $P^{TRQ}$
- $1 + t$
- Tariff Quota
- $W = 1$
- World Price
- Rent
- Tariff

Graphical representation showing the welfare analysis under TRQ with different price points and quantities.
The tariff quota limits imports at the in-quota tariff to $Q^{TRQ}$ and the domestic market clears at $P^{TRQ}$. This results in a further reduction in welfare: domestic consumers' surplus is reduced to the triangle labeled A and tariff revenue declines to the rectangle labeled ‘Tariff’. The area ‘Rent’ represents the arbitrage profits from the opportunity to import $Q^{TRQ}$ units at the cost of $1+t$, while the domestic market value is $P^{TRQ}$. These tariff quota rents are the result of rationing $Q^{TRQ}$ units of supply over $Q'$ units of demand. Rents are neither good nor bad, they simply exist as the result of rationing the opportunity to legally bridge the gap between domestic and world prices. Because there is no clear title to these rents they are a common property resource and, as such, their mere existence can stimulate wasteful rent-seeking behavior. How the rights to the rent are distributed largely determines the pattern and volume of exchange.

The result, $A+Rent$, is the best outcome possible given that imports are limited to $Q^{TRQ}$. Drawing the figure in this typical textbook manner assumes that a market mechanism is employed to ration $Q^{TRQ}$ units of supply over $Q'$ units of demand. That is, it assumes the rent is perfectly and automatically allocated in the best possible manner. This is a very strong assumption, if it does not hold, the outcome can be far inferior to the auction allocation.\(^7\)

### Auction Allocation

If one wishes to maximize welfare, given a tariff quota constraint on imports, one needs to devise some means of administering the tariff quota so that extramarginal suppliers are excluded: this is the beneficial discriminatory role prices play in free markets. The quota rent and the distorted incentive it transmits induces inefficiencies: demand inframarginal to W but extramarginal to P will enter the market as will supply inframarginal to P but extramarginal to W. If one could just get rid of the quota rent only inframarginal traders would enter the market and welfare would be maximized. An auction neutralizes quota rents.\(^8\) The opportunity to buy something worth W and sell it for P is worth $R = P-W$. Inframarginal traders will bid R for the opportunity; extramarginal firms bid less than R; and the required discrimination is realized.

There are several kinds of auctions. Perhaps the most familiar is the open-call, ascending bid auction: the auctioneer starts the bidding at some minimum value and bidders compete by bidding progressively higher prices. A sale is concluded when a bid remains unchallenged: the item is sold at the highest bid to the highest bidder. Such auctions work well for works of art and used cars. But often it is not feasible or worthwhile to bring all potential bidders together in one place for an open-

---

\(^7\) This point is generally overlooked in the literature. Vousden (1990):60-83, for example, devotes a full chapter to quotas without considering how quota rights are allocated.

\(^8\) There is a large and growing literature on auctions. The late William Vickrey (1961) initiated the economic literature on auctions. Milgrom (1989) is a readable introduction to the auction literature; Ashenfelter (1989) provides good illustrations. Campbell (1995) is a lucid guide to the more important proofs and places auctions and bidding in a demand-revelation context. Recent surveys and relevant applications include: McAfee and McMillan (1987); Wolfstetter (1996); and Das and Sundaram (1997).
call auction. The U.S. Treasury, for example, sells U.S. government debt obligations (bonds, notes and bills) through sealed-bid auctions. The Treasury gives advance notice of the amount and maturity of debt it will offer and the deadline for submitting sealed bids.

There are two kinds of sealed-bid auction: multiple-price and uniform-price. Suppose all prospective bidders for quota rights are asked to submit sealed bids specifying the amount of quota and the price per unit they are willing to pay. Once all bids are submitted the quota administrator opens them, ranks them in descending order by bid, and allocates bidders the amount requested until the quota is exhausted. In a multiple-price, sealed-bid auction each successful bidder is charged the price he or she bid for the quota. This results in price discrimination, different bidders are charged different prices for identical rights. Under a uniform-price rule, all winning bidders are charged the same price, specifically, they are charged the highest losing bid--this is the marginal bid and the best approximation to the market clearing price of a competitive market.9

Economists have maintained that uniform-price auctions are superior to multiple-price auctions for Treasury instruments. A series of bid manipulation scandals in 1990 and 1991 revived interest in the topic.10 The primary criticism of multiple-price auctions is that bidders have a much greater incentive to discover what other bidders are planning to bid than under uniform-pricing. In a uniform-price auction all that matters is whether one bids more or less than the highest losing bid. Under multiple-price rules, winners are, in effect, penalized for bidding more than the highest losing bid, thus there is an incentive to invest resources in discovering competitors’ intentions. Such additional investment in market intelligence is a waste of resources from the standpoint of allocative efficiency. Both auctions result in the same segregation of inframarginal and extramarginal bidders, but the uniform price auction requires less effort.

In a TRQ auction consumers would bid, at the margin, \( R = P - (1+t) \): the difference between the domestic price \( P \) (given imports of \( Q \)) and the world price plus tariff, \( 1+t \). If all winning bids are charged the marginal winning bid (uniform price auction) then auction revenue is shaded rectangle ‘rent.’ The consumers who obtain the quota rights are those with a willingness to pay of at least \( P \). These consumers realize a consumer surplus equal to the area \( A \). The domestic economy realizes gains from trade equal to the auction revenue plus \( A \). This allocation is identical to the allocation which would result from the tariff-equivalent of a tariff quota (given market conditions): \( t^* = t + R \).

Suppose we employ some alternative allocation method, for example, first come first served (FCFS). Under FCFS any consumer with a willingness to pay greater than \( W \) will attempt to import the instant the tariff quota season opens: the quota rent attracts all importers inframarginal to \( W \), many of whom are extramarginal to \( W+R \). To perform as well as an auction, alternative allocation methods must discriminate perfectly between agents inframarginal to \( W+R \) from those extramarginal to \( W+R \).

---

Why so few TRQ auctions?

Auctions are employed to administer only six percent of all TRQs notified to the WTO.\(^\text{11}\) If auctions are so wonderful, why are they not more commonly used? The economic answer is that auctions are most likely to outperform other rationing methods only if the market for the controlled product is sufficiently liquid. As a market becomes less liquid, its capacity to function as a price discovery mechanism deteriorates. Illiquidity has resulted in the demise of many commodity futures contracts.\(^\text{12}\) Those commodities for which active futures or cash markets exist are excellent candidates for quota auctions. If illiquidity diminishes the relative efficiency of an auction then other methods, such as license on demand, might be preferable. Related to the liquidity of the market is the number of active traders. Research on double auctions shows that too few traders can also diminish market efficiency. However, it does not require more than a few entrants to realize close to 100 percent efficiency. Several TRQ allocation methods specify a maximum market share that can be obtained by any single trader. Such rules can limit small numbers risk.\(^\text{13}\)

There is also a political explanation for the relative lack of auctions. Auctions are markets; and markets can be hard to control. If the government administering the TRQ has strong preferences about who should and should not gain quota rights then it will choose not to ration by auction. Similarly, if the government prefers to transfer quota rents to a select group rather than collecting the rents as auction revenue, it will not auction. As shown below, discretionary allocation --State Trading and Producer Group--provides the importing government or industry the greatest control over the distribution of rent and of trade.

### 4.3 Quasi-Market Methods

Several allocation methods are mixtures of market processes and random processes. This section examines First Come First Served (FCFS) and License on Demand (LD) as well as Historical allocation and Auctions as special cases. It considers, in turn, the risk these methods can pose of quota underfill and of biased distribution of trade.

A: Underfill

License on demand

License on demand allocation generally operates in the following manner. Before commencement of the quota period, potential importers are invited to apply for import licenses. Applicants specify the quantity they desire to import. Call the \(i\)th applicant’s demand \(q_i\), and call the sum of all import

\(^{11}\) Many notified auctions only occur when import demand is expected to exceed the in-quota volume, so even six percent is an overestimate.


\(^{13}\) Workshop on Double Auctions (1991), see also Tomek (1980).
application requests $Q^* = \Sigma q_i^*$. If domestic demand is sufficient the quota binds: $Q^* > Q^{TRQ} = \lambda Q^*$. To ration license supply among license demand, application quantities are reduced proportionately by the factor $\lambda < 1$. If one applies for $q^*$ units and the quota is binding, then a license is granted for $\lambda q^*$ units of imports at the in-quota tariff. Many countries also specify some minimum license quantum, so that the allocation rule reads: $\lambda q^*$ units but no less than $\gamma$ units. This minimum quantum rule can prevent the minimum shipload problem discussed below and in Appendix B.

The proportional reduction of requests complicates importing. First, if one accurately states one’s desired import volume and the quota is binding one will get less quota than desired. This creates the incentive to overstate one’s license request. If all applicants overstate by the same proportion the final allocation of shares is not affected, but if the degree of overstatement is not uniform, shares will be reallocated. In any event, planning is made unnecessarily difficult by this uncertainty.

Second, the physical carriage of goods is generally conducted in discrete units such as full truck or full container loads. Unit transport costs are generally minimized when loads are full as there are substantial fixed costs in shipping. Typically a license request will be for a multiple of full loads, say for $L^*$ containers. With proportional reduction only $\lambda L^*$ containers will be licensed for import at the in-quota tariff. Except in the rare case that $\lambda L^*$ results in an integer, proportional reduction will cause a remainder of one less-than-truckload shipment for each license application. Assuming for the moment that licenses cannot be divided and traded following their initial allocation, shippers holding rights to a less-than-truckload shipment face what amounts to a rounding decision. Above some fraction of a full load, say 80% it is still profitable to ship so the shipment is made; below the critical load of 80% the shipment is not made and the less-than-truckload portion of the license is not used. These unused remainders can result in underfill. [For an elaboration see Appendix B.]

**First come, first served: FCFS**

The standard FCFS allocation allows importation at the in-quota tariff until the quota has been filled. If domestic demand is sufficient to fill the quota then there is one individual shipment that fills the quota; that is, there is one last shipment that will be fully or partially over-quota. Call this last in-quota shipment the $n^\text{th}$ shipment and represent its volume as $q_n$. The tariff charged on the $n^\text{th}$ shipment is variable: the first $n-1$ shipments enter at the in-quota tariff; only a portion ($\alpha$) of the $n^\text{th}$ shipment enters at the in-quota rate with the over-quota balance $(1-\alpha)$ being charged the over-quota tariff; and for shipments after $n$, the over-quota tariff is applied.

Being the claimant of the $n^\text{th}$ shipment can be rather costly. Consider some of the alternatives. If it is costly to break the shipment - suppose it is a container or truckload - then the importer must choose whether to cease importing and route the whole shipment elsewhere or to pay the over-quota tariff on the proportion over-quota. If the shipment can be broken, then the over-quota portion can be shipped elsewhere (or forfeited at the border) leaving only the in-quota portion to be imported.

The system just described assumes that all customs agents have timely and accurate information on the level of quota utilization. Several countries have unified electronic monitoring systems and it is
possible to identify and inform the $n^{th}$ shipper prior to clearing customs. However, many countries do not have such systems. Information that the quota has been filled is disseminated to customs offices days, weeks or sometimes months later. Consequently it is possible to clear a shipment at the in-quota tariff and be notified later that payment of the tariff difference \([T-t]\) is due on each unit previously cleared. This additional charge will be levied only on shipments after the $n^{th}$ shipment, with the $n^{th}$ shipment liable for payment on \((1-\alpha)q_n\). If one discovers that one’s shipment is fully over-quota only after the fact, it is generally impossible to ‘un-import’ the shipment and avoid the extra tariff. If the over-quota tariff is significantly higher than the in-quota tariff the cost of being $n^{th}$ or later is very high. In the absence of timely information on quota fill, potential importers may avoid shipping if they believe that the quota is close to being filled and the risk of being caught over-quota is high. This rational risk-aversion can result in a lack of quota-fill, even when domestic prices exceed the in-quota landed value.

From the point of view of the government choosing to administer via FCFS there are incentives to delay notification. If collecting tariff revenue is a priority then sending the signal that the quota is not filled when in fact it is can entrap importers at the over-quota tariff. If protecting the domestic import competing industry is a priority, sending the signal that the quota is filled when in fact it is not will inhibit imports within the quota and yield the desired extra protection.

**FCFS Import Derby**

Another disadvantage of first-come, first-served allocation is that it can disrupt markets, as Johnson (1950:22) explains:

> The rush of products across our border at the beginning of each new tariff quota is a disquieting factor in domestic markets. The tariff quota on cattle illustrates this point very well. Before World War II the Canadian tariff quota was on a quarterly basis. At the beginning of each quarter Canadian farmers would rush their cattle into American markets, particularly at St. Paul, in order to take advantage of the lower tariff of 1.5 cents per pound instead of the above quota tariff of 3.0 cents per pound. The result was that the prices of cattle, particularly certain grades, on the St. Paul market were depressed with loss to both Canadian and American farmers. The quota added to price instability rather than reduced it. This example shows the competition for quota rents among firms from a single country. The potential for conflict is amplified when first-come, first-served competition is open to firms from several countries. There are three kinds of costs which result from the rush to the border: 1) an unnecessary dip in domestic prices; 2) unnecessary domestic storage costs; and 3) other, unnecessary rent-seeking costs induced by the existence of a common resource.

**B: Distribution of trade**

The risk of a biased distribution of trade from TRQ administration requires an examination of the supply side of the rationing problem. Figure 3 focuses on the supply side of the TRQ constrained market. To reduce clutter assume that the in-quota tariff is zero; that the over-quota tariff is
prohibitive, given market conditions; and that the domestic market clears at the price, P. The unit quota rent is \( R = P - W \). The figure plots an upward sloping supply curve ‘\( S' \); the supply elicited at price P is normalized to 1.

The existence of quota rents will induce suppliers inframarginal to P to enter the market. As shown in Figure 2 for the demand side, if the quota rights were auctioned, the quota rent would be absorbed in auction revenue, leaving a net import price of \( W = P - R \). Thus an auction only attracts supply inframarginal to W, efficiently solving the rationing problem.

From the point of view of welfare analysis and of GATT Article XIII it does not matter which firms or countries inframarginal to W gain access within the TRQ. Random displacement of inframarginal supply by other sources of inframarginal supply is not a problem. Displaced inframarginal supply can export to other markets at the world price. If rents are not fully absorbed through auction, tariff or other means, supply extramarginal to W will have an incentive to enter the market. This risks displacing inframarginal suppliers.

The area under the supply curve represents payments to factors employed to produce the traded product. Suppliers extramarginal to W must spend more than W per unit to employ labor, capital and other productive resources to produce a unit of output with a market value of only W. Rather than adding value, extramarginal production destroys value. This economically wasteful misallocation of resources is represented by the shaded triangle: Extramarginal Factor Use.

License on Demand

License on demand allocation may be thought of as a form of lottery. All firms inframarginal to P have an incentive to enter the quota lottery. The sum of applications, as shown above in the analysis of underfill, exceeds the supply of quota: \( Q^* = 1 > Q^{TRQ} = \lambda \). Each applicant wins a pro-rata share of the global quota: that is, \( \lambda q^* \) units. The effective supply curve for this uniform allocation of quota rights is \( S^0 \). The proportion of quota rights granted to suppliers inframarginal to W is \( \beta \). Given the assumptions made, LD allocation causes the displacement of \( \lambda(1-\beta) \) inframarginal suppliers by \( \lambda(1-\beta) \) extramarginal suppliers. This expected distribution of trade differs from the tariff equivalent counterfactual distribution, which consists exclusively of inframarginal suppliers. The welfare loss from LD allocation is shown by the shaded triangle below \( S^0 \).

This welfare loss can be interpreted as an indicator of how the expected distribution of trade differs from the tariff equivalent counterfactual distribution of trade. The maximum surplus possible is realized by market (auction) allocation. Reductions in welfare from this maximum occur because of the displacement of inframarginal suppliers by extramarginal suppliers. The greater the degree of displacement, the greater is the reduction in welfare, and the greater the difference in the distribution of trade from the counterfactual standard.
Figure 3 Quasi-Market Allocation
FCFS

To analyze FCFS allocation some assumptions must be made about the relationship or correlation between a supplier’s willingness to supply and its place in the FCFS queue. If a lower cost supplier always places ahead of every higher cost supplier, then the effective quota supply curve is identical to the original supply curve ($S$) in the interval $(0, \lambda)$: $S^\cdot$. However, because all suppliers can sell on the international market, $W$ becomes the lower bound on the willingness to supply the TRQ market for suppliers inframarginal to $W$. If each inframarginal supplier is equally likely to supply the TRQ market, it amounts to a random drawing from among the set of inframarginal suppliers (an operational test of non-discrimination). Such a random drawing yields the effective quota supply curve: $S^\ast$. Such a selection process also achieves efficient discrimination between inframarginal and extramarginal suppliers and produces an expected distribution of trade equivalent to an auction allocation. As a practical matter, however, it is difficult to conceive how such an efficient sorting of applicants could be achieved outside of using an auction.14

These two FCFS allocations ($S^\cdot$ and $S^\ast$) assume that there is a perfect correlation between lower cost and place in the queue. This is a strong assumption, but it is not implausible to assume that there might be some imperfect correlation. The License on Demand case above is equivalent to a FCFS allocation when there is zero correlation between cost and rank: a random selection of $\bar{e}$ units from a population uniformly distributed over the interval $(C,P)$ has the expected distribution represented by the curve $S^0$. A FCFS process with a positive but imperfect correlation will generate an expected allocation of quota rights among suppliers which may be represented by a supply curve drawn within the area bounded by $S^\ast$ and $S^0$ over the range $(0, \lambda)$.15 In sum, if a FCFS process generates a positive yet imperfect correlation between low cost and place in the queue, then there is some expected inframarginal displacement. The lower the correlation, the greater the expected displacement.

Finally, consider the perverse case where higher cost suppliers tend to place before lower cost suppliers. This is not as implausible as it might seem: one could argue that rent-laden markets are the only markets where extramarginal producers can hope to cover their factor payments, so they may specialize in getting to the front of tariff quota queues. [The necessary assumption is that loss aversion provides a greater incentive than profit maximization.] When low cost is perfectly and inversely correlated with place in queue (i.e., $\rho = -1$) the expected outcome may be represented by the supply curve marked $S^-$. This is the worst possible outcome, from a welfare and an Article XIII perspective: it results in the lowest realization of producers’ surplus possible, given that the quota fills; all inframarginal suppliers are displaced by the ‘most’ extramarginal suppliers. The triangle above $S^-$ is the producers’ surplus realized by the most extramarginal suppliers: it is equivalent to the shaded triangle in the upper right corner of the graph. The area below $S^-$ in the range $(0, \lambda)$

\[14\] Vickrey (1961) proves that the only allocative mechanism which will guarantee this result is a uniform (second) price auction, now commonly known as a ‘Vickrey auction.’

\[15\] The supply curve is a cumulative distribution of willingness to accept / to supply; it is a piecewise continuous, monotonically non-decreasing function. Under mild assumptions it can be shown that as the value of Pearson’s $\rho$ or Kendall’s $\tau$ ranges from zero to one, the corresponding supply curve rotates clockwise from $S^0$ to $S^\ast$. 
represents wasted resources and constitutes a welfare loss. Moreover, as established above, it indicates a biased distribution of trade relative to the counterfactual norm.

To summarize: this section has, so far, surveyed the range of welfare and bias effects possible under Auction, License on Demand, and FCFS allocations. It has shown that Auction and License on Demand (given the assumptions) are special cases of a general FCFS process. An Auction is equivalent to FCFS with perfect correlation between low cost and place in queue; but such perfect sorting is not possible without recourse to an auction. License on Demand is equivalent to a random FCFS process, with a zero correlation between cost and place in queue. The next section shows how historical allocation can also be characterized as a variant of FCFS allocation.

### Historical Allocation

Historical allocation can be viewed as an extreme variant of a FCFS process. FCFS and LD are annual lotteries. Each quota season there is a new drawing of applicants inframarginal to P. After several draws, the average realization will converge to its expected value. Historical allocation, in contrast, is essentially a one-time-only drawing. The historical base is infrequently revised: one particular realization is sustained for many years and remains invariant to changing market conditions. For example, exporter shares of the quota for US sugar imports were first allocated in 1934 on the basis of trade volumes in 1931-33. Save for wartime controls, the allocation was essentially unchanged until 1948. Legislation in 1948 and 1956 made minor adjustments to the shares of the two major suppliers: Cuba and the Philippines. The trade embargo imposed on Cuba after the Cuban Revolution forced a reassignment of the large Cuban share in 1961. It was formally reallocated in 1965 to countries, other than the Philippines, in proportion to their shares of the trade in 1963 and 1964. This allocation remained until 1974 when the 1948 quota was not renewed: imports were no longer limited by quota. A binding quota was re-imposed in 1982 on the basis of trade shares during 1975-81; this allocation was transferred unaltered into a tariff quota in 1995 and remains in effect. Each major change was prompted by an economic or political shock that, in each case, altered the structure of the sugar market. Despite this, the allocation of shares was based on the pattern of trade prevailing before the change. On average there has been about 15 years between major reallocations. The US Sugar TRQ is considered below in a case study where the extent of misallocation is calculated.

### Article XIII and Historical Allocation

The discussion of Article XIII in Section 1 emphasized its advocacy of non-discrimination. The analysis in section 2 has established that auctions, if practicable, are the best way to assure a non-discriminatory distribution of trade in quota constrained markets. FCFS and License on Demand, are inferior to auctions, and will generally result in a biased distribution, and historical allocation amplifies the bias. Despite these predictable biases, all four methods are consistent with Article XIII. It is in XIII 2(c) and 2(d), the sub-paragraphs on supplier quotas, where the contradiction between advocacy of non-discrimination and tolerance (if not advocacy) of discrimination is most clearly displayed.
GATT Article XIII, 2(c) states

Except in the case of supplier tariff quotas import licenses shall not require that the imported product originate from a particular country or source.

Supplier tariff quotas, also known as ‘allocated tariff quotas,’ are tariff quotas that are allocated to supplying countries, rather than to domestic importers or traders. The particular country or exporting firm and country is specified by the assignment of the tariff quota rights. XIII, 2(c) essentially states that importing countries can employ TRQ rights as a GATT consistent means of discrimination.

GATT Article XIII, 2(d) states

In cases in which a quota is allocated among supplying countries, the contracting party applying the restrictions may seek agreement with respect to the allocation of shares in the quota with all other contracting parties having a substantial interest in supplying the product concerned. In cases in which this method is not reasonably practicable, the contracting party concerned shall allot to contracting parties having a substantial interest in supplying the product shares based upon the proportions, supplied by such contracting parties during a previous representative period, of the total quantity or value of imports of the product, due account being taken of any special factors which may have affected or may be affecting the trade in the product.

The two italicized phrases (here, not in the original) have been the subject of further definition by the GATT in a series of interpretative notes to Article XIII. The convention has been to use an average of the three years prior to the imposition of a restriction as the representative period. Several disputes have arisen over base periods during which there were other restrictions on trade. The GATT recommends that shares be allotted according to the trade shares “which would correspond to what could reasonably have been expected in the absence of restrictions.” Once again, this is the free trade counterfactual distribution of trade, the operational equivalent of non-discrimination.

With regard to the meaning of special factors:

The term ‘special factors’ as used in Article 22* includes among other factors the following changes, as between the various foreign producers, which may have occurred since the representative period:

1. Changes in relative productive efficiency;
2. The existence of new or additional ability to export; and
3. Reduced ability to export.

[*Article 22 refers to the article of the Havana Charter corresponding to GATT Article XIII: 2(d) and 4.]

Thus XIII: 2(c) and 2(d) instruct member governments that they are allowed to transfer TRQ rights
Skully: TRQ Administration

to incumbent exporters, but they should do so in such a way as to approximate the free trade counterfactual distribution of trade. This is not a simple task. The passage above elucidating the term ‘special factors’ gives the impression that exporter shares can be reallocated in line with changing economic conditions, presumably without compensation. However, once exporters are vested with quota rights they tend to become upset when there is the least suggestion of taking them back or transferring them to a competitor. The author is unaware of a case where this kind of reallocation has occurred in accordance with Article XIII. The lack of such reallocations is hardly surprising. The primary reason an importing government chooses to allocate ‘supplier quota’ is to appease suppliers harmed by the imposition of a quota. The U.S. tobacco and sugar TRQs are examples. The quota rights are non-transferable; and the product delivered in-quota must be the domestic product of the exporter. Such restrictions are the cause of biased trade shares and often (for tobacco) of quota underfill. The corollary to the last statement is that removing such restrictions can remedy, or at least substantially reduce, the risk of bias and underfill.

Resale Markets in TRQ Rights

When TRQ rights are not transferable, one must export to realize the (compensatory) rent. Thus the distribution of rent and the distribution of trade are identical. Allowing the transfer of TRQ rights liberates the distribution of trade from the distribution of rent. From the perspective of the WTO, the only relevant consideration is the distribution of trade, not the distribution of rent. This means, for the purpose of evaluating alternative methods of TRQ administration on the basis of non-discrimination, one should ignore the redistribution of rent. Of course, it is the redistribution of rent that drives much of the politics of quota administration. As Gardner (1983) has demonstrated, the transfer efficiency of commodity programs (including schemes analogous to supplier quotas) is relatively low. It is far better to compensate legitimate rent claimant by direct monetary transfers than through discrimination and manipulating markets.

The result above holds only if there is no resale of quota rights. Resale - scalping, arbitrage - can reclaim much of the deadweight loss caused by extramarginal suppliers displacing inframarginal suppliers. If there are no transactions or information costs, and all agents are rational profit-maximizers, then extramarginal quota holders will sell their quota rights to inframarginal suppliers discriminated against by quasi-market allocation processes. These trades, in a perfect market, occur at the price R, the marginal auction bid defined earlier. Extramarginal quota holders value quota at R or less, while inframarginal suppliers value the quota at R or more.

When resale is allowed the final distribution of rent will differ from auction allocation only in that auction revenue is captured by private traders rather than by the government. Suppliers inframarginal to W who receive quota rights keep them; suppliers extramarginal to W who gain quota rights in the primary allocation sell them. Essentially the auction revenue/rent is redistributed from the government or auction authority to private agents. Some countries, states and provinces allow the resale of quotas and quota rights. In fact the government often organizes and supports the market...
exchange institutions. Quota or rights trade has been allowed or devised to reduce air pollution and to prevent over-fishing.

**US Sugar TRQ – An example of Historical Allocation**

The U.S. sugar tariff quota is allocated to exporting countries on the basis of their ‘olympic average’ market shares of US sugar imports in the period 1975-81. This was a period of exceptionally high world sugar prices. So high, in fact, that in 1975 the United States removed the quantitative import restriction that had been in place since 1934. During several months of the base period the world price of sugar exceeded 30¢ per pound. At 30¢ virtually everybody is an inframarginal sugar supplier. The market shares of US imports during the base period 1975-81 included some suppliers who happened to be inframarginal for a few months, but are extramarginal under ordinary market conditions.

Skully (1998) examines the pattern of imports for ‘quota exempt re-export sugar.’ Raw sugar may be imported to the United States outside of the quota if it is refined and re-exported within a 90 period. This trade is not distorted by tariffs or quotas (save for the embargo on Cuba) and can be used as an estimate of the free trade counterfactual distribution of trade. This distribution is contrasted with the allocation of TRQ shares in the pie charts. The quota exempt distribution of trade is dominated by low cost sugar producers located relatively close to U.S. refining centers in Gulf and Atlantic ports.

The requirement that sugar imported under the TRQ must be produced in the country allocated the quota rights amounts to an anti-scalping law. This restriction induces costly transactions. Taiwan, for example, has tariff quota rights for the export of about 24,000 short tons of sugar to the United States. Taiwan always fills its quota; however, this is the only sugar it exports. Taiwan’s domestic production does not satisfy its domestic demand. It imports sugar (usually from Australia or Thailand) to cover the difference, including an additional 24,000 tons to cover the domestic production exported to the United States. It would be more efficient for Taiwanese quota holders to charter a shipment of 24,000 tons of sugar from Queensland or Guatemala to the United States and simply pocket the arbitrage rents. Similarly, the Philippines, the third largest quota holder (13%), has recently been unable to cover its domestic needs from domestic production: in fact, it has a TRQ

---

16 Alston (1986) examines egg quota trading in Victoria; Canadian milk quotas have generated a large literature: e.g., Barichello (1996) and Chen and Meilke (1998); for dairy quota markets in the EU see: Oskam (1991) and Pennings and Meulenberg (1998); finally, Rucker, Thurman and Sumner (1995) is an excellent study of the restrictions on the transfer of flue-cured tobacco quota in North Carolina, it distinguishes the distribution of rent from the distribution of production.

17 The current TRQ was converted from a standard quota after Australia successfully challenged the U.S. quota in the GATT on the grounds that it violated Article XI. See: United States - Restrictions on Imports of Sugar. BISD 36S/331. June 22, 1989. Establishment of the TRQ resolved the dispute. Also of interest is the Nicaraguan case: United States - Imports of Sugar from Nicaragua. BISD 31S/67, 73. March 13, 1984. Nicaragua successfully argued that the United States violated Article XIII:2 when it unilaterally reduced Nicaragua’s share of the sugar quota.

18 A short ton is 2,000 pounds; a Metric ton is 1,000 kilos or 2,204.62 pounds.
Distribution of Trade for U.S. Sugar Imports

Quota Exempt Shares

- Guatemala
- Colombia
- Other Cent. Am. and Carib.
- Other

TRQ Shares
to limit sugar imports. To procure domestic sugar to fill its US tariff quota, the Philippine sugar authorities offered domestic mills 1.2 tons of imported raw sugar for every ton of domestic raw sugar delivered for export to the United States.

In both these examples the ability to resell tariff quota rights would improve international factor allocation. Because the revenue from resale or arbitrage could easily fund compensation (in the case of Taiwan) or fund investment (in the case of the Philippines) the resale would not have major domestic political repercussions. This is not the case for all current sugar TRQ holders. Several nations in the Caribbean region have both U.S. and EU preferences for sugar exports (e.g., Guyana, Barbados, St. Kitts-Nevis). Even with the quota rent income, sugar production is, at best, barely a viable economic activity. Because sugar production accounts for a large share of domestic employment, and sugar workers are well organized, the sale of quota rights would likely precipitate massive unemployment and/or a political crisis. In the absence of alternative employment opportunities, these governments might retain the quota rights to preserve the domestic status quo.

4.4 Discretionary Methods

Discretionary methods of TRQ administration are those which delegate the allocation process to a select group or organization. In the case of State Trading, import rights within the TRQ, and sometimes outside the TRQ, are granted to specialized government bureau. In the case of Producer Group administration, the import rights are granted to an organization that represents producer interests. How these organization choose to exercise TRQ rights is limited only by the discretion granted them by their governments, or, in the case of producer groups, by their membership.

Table 4 lists all the TRQs administered by discretionary methods. There were 22 State Trading TRQs and 9 Producer Group TRQs notified to the WTO for 1995 and 1996. A review of the list reveals that Thailand and South Korea account for the majority of state trading TRQs and all producer group TRQs. Most are for products that are trivial from the perspective of global agricultural trade: e.g., pine nuts, raw ginseng, onion seed, garlic, capsicum, sesame, potatoes. There are eight state trading TRQs which are potentially important: four are for rice: Indonesia, Japan, South Korea and the Philippines; the Japanese TRQs for wheat, barley and dairy are important simply because of the size of the Japanese market; finally, there is the Thai tobacco TRQ. None of the Producer Group TRQs is particularly important. While few of these TRQs are important in a global perspective, they can still be a potential source of conflict among interested suppliers. Thus they should not be dismissed casually.

For reasons discussed below and above in Section 3, both State Traders and producer groups tend to fill their quotas if there is sufficient domestic demand. So, quota fill risk is generally not a problem under discretionary administration; the problem lies in the distribution of trade. Because the import sourcing decision is made by public or quasi-public officials and not by private traders, there is a multitude of factors divorced from commercial considerations which may determine market shares. That is, political considerations play a major role. It would not be fruitful to generalize beyond the few generalizations already put forth, because, unlike quasi-market methods which can be reduced
Table 4

State Trading Organizations
(n=43)

![Bar chart for State Trading Organizations]

Producer Groups (n=17)

![Bar chart for Producer Groups]
### Table 4 cont.

#### STATE TRADING

<table>
<thead>
<tr>
<th>Country</th>
<th>Commodity</th>
<th>Fill rate 1995</th>
<th>Fill rate 1996</th>
</tr>
</thead>
<tbody>
<tr>
<td>canada</td>
<td>butter</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>indonesia</td>
<td>rice</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>japan</td>
<td>dairy 12</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>wheat</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>barley</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>rice</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>korea</td>
<td>potatoes</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>onions</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>garlic</td>
<td>77</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>capsicum</td>
<td>99</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>beans dry</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>ginger</td>
<td>96</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>rice</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>buckwheat</td>
<td>100</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>groundnuts</td>
<td>100</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>sesame</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>philippines</td>
<td>rice</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>thailand</td>
<td>milk</td>
<td>0</td>
<td>PG*</td>
</tr>
<tr>
<td></td>
<td>potatoes</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>garlic</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>coffee</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>tobacco</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

#### PRODUCER GROUP

<table>
<thead>
<tr>
<th>Country</th>
<th>Commodity</th>
<th>Fill rate 1995</th>
<th>Fill rate 1996</th>
</tr>
</thead>
<tbody>
<tr>
<td>korea</td>
<td>pine nuts</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>citrus</td>
<td>98</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>raw ginseng</td>
<td>100</td>
<td>1</td>
</tr>
<tr>
<td>thailand</td>
<td>milk</td>
<td>STE*</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>onions</td>
<td>0</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>onion seed</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>soy oil</td>
<td>92</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>palm oil</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>sugar</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* Thailand switched its milk TRQ from STE to PG in 1996
plausibly to algorithms, each discretionary institution has its own methods, or, perhaps more accurately, non-methods. The best one can do, at least at this initial stage of TRQ analysis, is to examine them individually. Thus, the balance of this section consists of a case study of how the Japanese Food Agency, a State Trading Organization, administers the Japanese TRQ for Wheat. One case study cannot pretend to be a representative sample. The case was selected because of the availability of data and secondary information as well as its importance in U.S. agricultural trade diplomacy.

**Wheat Imports by the Japanese Food Agency: An Example of Discretionary Allocation**

Japanese grain and oilseed imports are not uniformly regulated. Corn, soybeans and most other oilseeds are imported with relatively few restrictions; rice, until the Uruguay Round, was simply not imported; and wheat imports were and continue to be controlled by the Japanese Food Agency [JFA]. The JFA has been notified to the WTO as a State Trading Enterprise [STE].

The JFA is one of the world’s largest importers of wheat. Consequently, it has been the subject of considerable study. Unfortunately, much of the economic analysis of the JFA’s import procurement decisions has been predicated on erroneous as if assumptions about the JFA’s objective function. This section draws upon one study that presents a persuasive argument about what the JFA’s objectives are. Alston, Carter and Jarvis (1990) argue that the JFA operates a discriminatory quota scheme for wheat and that the Japanese Livestock Industry Promotion Corporation [LIPC] operates a similar regime for beef imports. Each year the JFA, in consultation with the domestic milling industry and MITI (the Ministry of International Trade and Industry), determines domestic consumption targets for wheat and wheat products, estimates expected domestic wheat production, and determines aggregate import needs. The aggregate quota is procured by subcontracting to a pool of about 30 trading firms. The contracts specify the variety of wheat required as well as a designation: food wheat or feed wheat. As Alston et al. show, the food/feed distinction is not particularly meaningful. Rather it allows an additional degree of freedom in determining supplier market shares. At the time Alston et al. drafted their article only Australia and the United States were permitted to supply feed wheat; Canada was excluded as a feed wheat source.

The essence of the Alston, et al. analysis is that the JFA’s allocation of supplier shares for wheat is analogous to a set of Voluntary Export Restraints [VER]. A VER shares quota rents with the exporting country that could be fully captured by the importing country. By the standard of domestic welfare maximization (the assumption employed in many studies of the JFA) a VER is clearly inferior to a global quota or its tariff equivalent. If an importing country chooses a VER it reveals that the welfare of the preferred supplier(s) is important to the importing country government.

...the government in the importing country attaches greater value to the appeasement of the foreign interests than it does to the efficiency costs of not using an otherwise economically superior instrument. Thus we can explain the Japanese government’s use of managed import quotas for beef and wheat, but to do so we must allow for the political influence of import suppliers. [p. 200]
Alston et al. base their conclusion on the observed imports of the JFA and the LIPC through 1989 (when their article was drafted) as well as supplementary evidence from interviews with principals in the trade. A telling piece of evidence is that the supplier import shares are too stable: a market driven allocation would show more variance than what is observed. As evidence of a JFA bias for U.S. wheat and against Canadian wheat, they note that Japanese milling firms consider No. 1 Canadian Western Red Spring wheat [CWRS] the most preferred variety. Mills always request more No. 1 CWRS in their annual consultation with the JFA, but the share procured remains, in their view, too low. On the basis of this evidence, Alston et al. conclude that a liberalization of the Japanese wheat market, for example, converting the JFA’s discretion over imports into a tariff equivalent or auctioned TRQ, would result in a transfer of market share from the United States to Canada. Moreover, they argue that Japanese wheat consumption is already relatively high in per capita terms and tariff reductions are not likely to lead to a dramatic increase in the volume of imports. Consequently, liberalization could lead to an absolute decline in the volume of US wheat exports to Japan. [The substitution effect would dominate the expansion effect. Alston et al., p. 210.]

There is now a decade of additional trade data to examine how well the Alston, Carter and Jarvis thesis holds. Figure 4 shows the market shares of the three suppliers of wheat to the JFA. [n.b., these are the only suppliers.] Shares are shown for food, feed, and total (combined) wheat. Basic descriptive statistics are included in the tables. The shares are plotted as values between 0 and 1 using a logarithmic scale. The annual shares are very stable. The only change is a minor increase in the Australian share of the food wheat market. The increase has been very gradual and the stable rate of increase is indicated by the constant slope of the log-share line. The Australian expansion in the food wheat market has been mirrored by the allowance of Canadian ‘feed’ wheat imports: a minor innovation since Alston et al. (1990).19

Arnade and Gehlhar (1999) analyse monthly wheat import data of 43 countries, accounting for over 90% of world wheat trade, for evidence of importer market power. In their discussion of importer sourcing behavior during 1962-1995 they note:

Most importers have shifted suppliers somewhat. Japan however is a clear exception; its shares have been stable. It sources nearly all of its wheat from the United States, Canada, and Australia in the fixed proportions of 55%, 27% and 18% respectively.

Arnade and Gehlhar also examine the monthly trade data for the frequency of multiple suppliers.
Table 5

JAPANESE FOOD AGENCY WHEAT IMPORTS: 1982-1997

### Food wheat

<table>
<thead>
<tr>
<th></th>
<th>Canada</th>
<th>U.S.</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.29</td>
<td>0.60</td>
<td>0.11</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.05</td>
<td>0.02</td>
<td>0.20</td>
</tr>
<tr>
<td>Coef. var.</td>
<td>0.02</td>
<td>0.47</td>
<td>0.52</td>
</tr>
</tbody>
</table>

### Feed wheat

<table>
<thead>
<tr>
<th></th>
<th>Canada</th>
<th>U.S.</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.02</td>
<td>0.47</td>
<td>0.52</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.99</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>Coef. var.</td>
<td>0.23</td>
<td>0.58</td>
<td>0.19</td>
</tr>
</tbody>
</table>

### Food and feed wheat

<table>
<thead>
<tr>
<th></th>
<th>Canada</th>
<th>U.S.</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>0.04</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>Coef. var.</td>
<td>0.04</td>
<td>0.02</td>
<td>0.06</td>
</tr>
</tbody>
</table>
They argue that a high frequency of single source observations may represent a series of corner solutions by a very price sensitive importer. One buys only only from the least cost supplier, whoever it is that month. Conversely a high frequency of multiple suppliers may reflect an inelastic demand for particular supplier product characteristics, or possibly a concern for a diversified supply. A high frequency of multiple suppliers is also consistent with the VER hypothesis of Alston, et al. Arnade and Gehlhar find that Japan has imported from its three regular suppliers 95% of the time; in only 4% of the months observed did it procure from only two of the three suppliers, and less than 1% of the time from only one supplier. With the exception of South Korea, no other country has a pattern of trade remotely similar to Japan. The persistence and stability of its pattern of imports is consistent with the VER hypothesis.

Mao et al. (1997) is a recent econometric study of the input substitution of wheat varieties by the Japanese milling industry. While not the primary thesis of the paper, the paper argues against the VER hypothesis advanced by Alston et al. and the present paper. They conclude that “the multiple classes and the multiple end-use characteristics of U.S. wheat enable the United States to maintain the largest market share in the Japanese wheat market.” (p. 12) This conclusion, however, is consistent with the VER hypothesis. The data Mao et al. analyze are generated by JFA sourcing decisions, they do not represent solutions to optimization problems by the Japanese milling industry as their analysis assumes. The fact that Japanese millers are constrained to mill such a large share of US wheat induces substitution and processing which would not occur were import procurement more closely matched to notional milling demand.

South Korea is mentioned above as the only country with a pattern of wheat imports remotely similar to Japan’s. There is a reason for this similarity: South Korea is a former State Trader in wheat. Until 1983, the Korean Flour Mills Industrial Association [KOFMIA], a state coordinated umbrella organization, held monopoly control over wheat imports.

The U.S. share of the Korean milling wheat import market was virtually 100 percent until 1983, when private importers were permitted to enter the market. ... U.S. milling wheat has been displaced in the Korean market primarily by Australian wheat, including Australian standard white (ASW), Australian Hard (AH) and Australian Soft (AS). Canadian Western Red Spring (CWRS) 13.5-percent protein wheat has begun to make inroads and will probably capture a larger share of the market [Raney and Morgan (1994:8)]

The graph plots the market shares of Korea’s major suppliers since 1980. The U.S. share since 1983, when KOFMIA was liberalized, has averaged 50%; the share is volatile, as one expects in a competitive market, and shows a secular decline. The Korean case cannot be taken as a perfect counterfactual for the liberalization of the JFA, but a decline in the U.S. share is likely, and an increase in its volatility is certain.

The JFA fills its import quota, but this is not surprising. In practice, State Trading poses a low risk of underfill. TRQs which implement Minimum Access commitments are generally carefully
scrutinized by potential suppliers. If there is market demand, quotas fill. The greatest risk is that the discretion over sourcing, being in the hands of public employees, is particularly vulnerable to political pressure. If some suppliers are capable of exerting more political pressure than others, they may gain a larger share of the quota constrained market. If such pressure alters the distribution of trade from its tariff equivalent counterfactual distribution, that is, based on supplier efficiency alone, it violates the principle of non-discrimination.

The United States has been especially aggressive in its efforts to gain access for its agricultural products in Japan, South Korea, and other Asian markets. In general, U.S. efforts have been bilateral. Free riding can dilute the gains from bilateral negotiation: other suppliers can gain market access if liberalization is applied multilaterally. The losses to free riding can be minimized if liberalization is de jure multilateral, but de facto bilateral.

US-Japanese economic controversies in the recent past have been most acute where GATT guidelines are vague or provide broad exemptions, ... Some of these have been overcome by means of bilateral negotiations between the two countries. The danger of such bilateral ‘deals,’ however, is their possibly unintended, but nonetheless real, impact on third, fourth, or twentieth parties. [Ehrenhaft (1990): 275]

Such bilateral ‘deals’ exist for several industries.

South East Asian rice producers, European telecommunications manufacturers, Canadian foresters, Argentine gaucheros, and Spanish tanners are all competitive with their American counterparts. Yet the American share of Japanese purchases of these products has been disproportionately high. [Oye (1992): 157]

These external effects on ‘third, fourth, or twentieth parties’ amount to a biased distribution of trade and a violation of the principle of non-discrimination. Placing greater discipline on State Trading Organizations has been targeted as a priority for the United States in the next WTO. When Administration officials speak of State Trading Organizations they invariably mean, for agriculture, the Canadian Wheat Board, the New Zealand Dairy Board or a handful of other export boards. The JFA, as a State Trading Organization, would also be subject to the same disciplines as the export boards. Constraining the JFA to procure in a less discriminatory manner would erode the U.S. share of the Japanese wheat market. Removing monopoly import power from the JFA would, like in case of KOFMIA, result in more volatile markets shares as well.

5 Conclusion

One of the issues to be resolved in the forthcoming round of WTO negotiations is what rules should govern TRQ administration. WTO negotiations on TRQ administration reduce to whether and how

---

20 If not, political pressure is applied, bilaterally or through the WTO. The recently resolved U.S.-Philippine dispute over the Philippines’ administration of its TRQ for pork follows this pattern: the quota eventually fills.
South Korean Wheat Imports

supplier market shares

Source: Korean Statistical Yearbook of Foreign Trade
to discipline the set of available administrative methods. Negotiations on TRQ liberalization reduce
to whether and how to bound the levels and time paths of the parameters which constitute a TRQ: {t, T, Q}. This analysis has limited itself to the former question -- the economics of TRQ administration -- in the light of one possible interpretation of GATT Article XIII. This interpretation emphasizes allocative efficiency as the fundamental normative principle of the WTO; thus, it places no weight on the distribution of TRQ rents. It recognizes that distribution matters, and demonstrates that equity and efficiency are largely separable and need not conflict. The exclusive focus on TRQ administration means that the analysis does not consider the equally important and related issue of TRQ liberalization, a topic that requires, and will undoubtedly attract, more attention.

**Implications for the Next Round**

The needed legal reform would be a proposal to delete from Article XIII the authority to impose county-specific QRs, forcing governments to rely instead on global quotas, administered by auctioning global licenses. Such a proposal would, no doubt, stimulate vigorous opposition from just about every quarter, and chances of GATT adopting such a proposal in the foreseeable future are quite poor. [Hudec, p. 206]

The economic analysis of this paper supports Hudec’s proposal: If TRQs cannot be abolished or converted to ad valorem tariffs, then the second best option is to make all TRQs global and auction global licenses. The author also agrees with Hudec that there is likely to be substantial political resistance to such a radical change in the rules of the game. But, auctions might provide sufficient revenue to compensate recalcitrant parties.

The economics of TRQ administration are clear if one considers only the allocative efficiency of the various methods of administration. On this basis, market methods dominate all others and therefore should be encouraged. Historical allocation and discretionary methods are the least desirable and therefore should be disciplined. These conclusions were presented above in section 4. It would be naive, however, to ignore distributional effects: who gets the rent does matter. In fact, in the realpolitik of trade policy, it matters far more than global allocative efficiency. The good news is that (given the existence of a TRQ) there is no necessary trade-off between allocative efficiency and rent distribution: the two are separable. This is true even without recourse to auction. One may also determine the distribution of rents by the initial assignment of quota rights. Then, by allowing the resale of allocated quota rights, market forces can match trading rights with inframarginal traders and permit the realization of most of the potential allocative efficiency in the product market. If rent entitlement claims cannot be satisfied by lump-sum transfers of auction proceeds then they may be satisfied as follows:

Distribute licenses to import within the TRQ in any manner the importing country prefers (thus distributing rents) provided that:

1] licenses be transferable and freely tradable to any trader.
2] there be financially significant penalties for license holders who do not import when there is unmet excess demand.

3] if licenses are distributed to STEs or producer groups, they must auction the licenses (with no restriction on participation) and distribute the proceeds in a non-trade-distorting manner (that is, decoupled: not contingent on the actions of the beneficiaries)

Discretion over the initial allocation may be used to satisfy priority claims of rent entitlement. Point 2 restrains the exercise of market power in the import market. In the market for quota rights, if imperfect competition is a threat to efficiency, then limits on the proportion of rights distributed to a single agent or alliance of agents could be imposed. There is a risk that the distribution of rents may have trade distorting effects, this holds whether rents are lump-sum or through quota resale. The trade distorting effects are likely second-order, but, if sufficiently large, would probably be liable to discipline under other GATT articles.
Appendix A
WTO and Auctioning TRQ rights

WTO legal arguments

Language in Articles II and VIII has been used to argue against initiating auction allocations of tariff quota rights. Given the economic interpretation of Article XIII presented in Section 2.1 neither argument alone is compelling and, when considered jointly, the logical basis for opposition to auctions is even weaker.

Article II  The argument is that auctioning import rights affords a greater degree of protection than allowed in the tariff schedule submitted to the WTO. From an economic perspective, this is simply not true. Shifting from an existing, non-auction method of tariff quota allocation to auctioning tariff quota rights does not change the degree of protection afforded. The potential volume of trade within the tariff quota does not change. There is only a change in distribution of quota rents, to which the WTO should be indifferent. Suppliers or importers are still allowed to import the same volume of goods at the in-quota tariff. The unit rent, the difference between the domestic price and world price less costs, is now absorbed as auction revenue instead of being captured by the holders of the tariff quota rights. Auctions only change the composition of the trade by discouraging extramarginal suppliers, thus resulting in an unbiased composition of trade as required by Article XIII. Auctions will not change the volume of trade.

Article VIII  The argument is that auctions constitute an extra fee on imports. Article VIII:1(a) requires that fees are only allowed to cover “the approximate cost of services rendered and shall not represent an indirect protection to domestic products or a taxation of imports or exports for fiscal purposes.” Allowing suppliers or imports to capture tariff quota rents is an unrequited transfer of a financial asset. Since the government created the restriction on trade, one may reasonably presume that it has a prior claim to the rents thereof. It may choose to transfer these rents, unrequited, and the cost of this gift equals the value of the rent so transferred. By auctioning tariff quota rights the government is merely charging the market value of the financial asset transferred. The GATT says nothing regarding who is entitled to rents (nor should it as these are neutral transfers). Former quota rent recipients have no actionable claim for damages because auctioning now requires them to pay (or bid) for assets which were once transferred to them free of charge.

The claim that auction revenue constitutes taxation for fiscal purposes is easily dismissed. Most tariff quotas were constructed to replace quotas disallowed by the Uruguay Round. These former quotas did not raise revenue, they provided domestic protection and transferred quota rents to select recipients. Neither the motive nor the effect of the original quotas was the generation of tax revenue for fiscal purposes; nor can this motive or effect be imputed to their tariff quota equivalents. Allocating tariff quotas by auction does generate a flow of revenue, but this revenue is requited by the transfer of rights of an equivalent expected value. Similarly, the auction of Treasury bonds by the
Skully: TRQ Administration

U.S. Treasury generates revenue which may be used for fiscal purposes; but as the sale of a bond is the exchange of an asset for funds, it is not considered a tax on the bond buyer.

*U.S. legal considerations*

Auctioning requires no new legislation: the executive power to auction import licenses is part of the U.S. Code. Section 1102 of the Trade Agreements Act of 1979 allows for the public auction of import licenses.

U.S. Code Title 19:
§ 2581. Auction of import licenses
(a) In general
   Notwithstanding any other provision of law, the President may sell import licenses at public auction under such terms and conditions as he deems appropriate. Regulations prescribed under this subsection shall, to the extent practicable and consistent with efficient and fair administration, insure against inequitable sharing of imports by a relatively small number of the larger importers.
(b) "Import license" defined
   For purposes of this section, the term "import license" means any documentation used to administer a quantitative restriction imposed or modified after July 26, 1979 under -
   (1) section 125, 203, 301, or 406 of the Trade Act of 1974 (19 U.S.C. 2135, 2253, 2411, or 2436),
   (2) the International Emergency Economic Powers Act (50 U.S.C. 1701-1706),
   (3) authority under the notes of the Harmonized Tariff Schedule of the United States, but not including any quantitative restriction imposed under section 22 of the Agricultural Adjustment Act of 1934 (7 U.S.C. 624),
   (4) the Trading With the Enemy Act (50 U.S.C. App. 1-44),
   (5) section 204 of the Agricultural Act of 1956 (7 U.S.C. 1854) other than for meat or meat products, or
   (6) any Act enacted explicitly for the purpose of implementing an international agreement to which the United States is a party, including such agreements relating to commodities, but not including any agreement relating to cheese or dairy products.

The author, a legal layman, infers that item b(3) ceased to have any referent following the conclusion of the UR agreement. All section 22 quantitative restrictions were converted to tariff quotas. Because the WTO does not consider tariff quotas to be quantitative restrictions (and if U.S. courts concur), it must follow that any TRQ, save for those explicitly excluded in b(5) and b(6) can auctioned at the discretion of the President.
Appendix B
License on Demand Underfill

One can make some inferences about the level of underfill when License on Demand is employed to allocate quota rights. First, suppose that all trade is conducted in some standardized container load of X kilos. Call Z the tariff quota volume expressed in units of container loads: $Z = \frac{Q^{TRQ}}{X}$. The total amount of less-than-container-load shipments will equal $N/2$ full container loads where $N$ is the number of applicants. This follows from the fact that the remainder is random and uniformly distributed over the range $[0,1)$ with an expected value of 1/2. With $N$ applicants the sum of remainders is $N/2$.

Call the critical load value $k$. Because the remainder values are uniformly distributed the expected value of remainders less than $k$ is $k/2$, and the expected value of remainders greater than or equal to $k$ is $(1-k)/2$. Similarly the number of remainders less than $k$ is $kN$ and the number of remainders greater than or equal to $k$ is $(1-k)N$. The quantity of underfill is the number of loads less than $k$ times the average quantity of such loads: $kN \times k/2 = \frac{1}{2}Nk^2$.

<table>
<thead>
<tr>
<th>Number of remainders</th>
<th>N</th>
<th>Average quantity</th>
<th>1/2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of 0 loads</td>
<td>kN</td>
<td>Average quantity</td>
<td>k/2</td>
</tr>
<tr>
<td>Number of 1 loads</td>
<td>(1-k)N</td>
<td>Average quantity</td>
<td>(1-k)/2</td>
</tr>
</tbody>
</table>

Expected underfill $\frac{1}{2}Nk^2 = kN \times k/2$

From the point of view of quota fill efficiency the relevant magnitude is the expected underfill as a proportion of the quota; define this value as $U = \frac{Nk^2}{2Z}$. Consider how changes in the components of this measure influence the value of $U$. An increase in $k$, the critical value of a shipment, increases underfill. The higher the critical value the less likely a less than full shipment will be made. An increase in the number of applicants, $N$, also increases underfill. This follows because the number of applicants determines the number of less than full load remainders. Finally, increasing $Z$ reduces underfill. This follows simply because $Z$ is in the denominator, clearly with $\frac{1}{2}Nk^2$ invariant to $Z$, an increase in $Z$ must reduce $U$.\(^{21}\)

\[ U = \frac{Nk^2}{2Z} : \quad \frac{dU}{dk} > 0; \quad \frac{dU}{dN} > 0; \quad \frac{dU}{dZ} < 0; \quad \frac{dU}{dX} > 0 \]

The value of $Z$ could increase for two reasons. First, the quota could simply be expanded. Second, innovations in shipping can change $Z$ even if $Q^{TRQ}$ is unchanged. If, for example, smaller sized

\(^{21}\) As long as $Z > N$, the value is invariant to the value of $Z$. When $Z < N$ the expected licence quantity is less than 1 and the assumption of a uniform distribution of the remainder over the unit interval no longer holds. When $Z < kN$ the average applicant will be granted a licence for less than one critical load and therefore will not choose to ship. The lower the value of $Z$ below $kN$, the more likely all applicants will choose not to ship and none of the quota will be used.
shipments become economically feasible the number of kilos per shipment decreases. Thus the number of shipments allowed by the quota increases: \( Z = \frac{Q_{TRQ}}{X} \), reducing \( X \) increases \( Z \). Finally, when the quota ceases to bind, \( Q_{TRQ} > Q^* \) implies \( Z > Z^* \), all license requests are granted, there are no remainders, and \( U \) collapses to zero.

Appendix C  
Economic literature on tariff quotas and related topics

The literature on tariff quotas may be divided into two genres. There is a descriptive literature, primarily from before WWII, which discusses problems in the administration of tariff quotas prior to GATT [e.g., Heuser (1939), Gordon (1941), and Delle Donne (1928).] From this literature some interesting facts emerge. First, tariff quotas have been around a long time. The earliest reference is to a 1839 Belgian tariff quota on the import of cast iron from Luxembourg. Second, tariff quotas were only widely employed after the First World War and particularly in 1931-32. However, the severity of the depression, with competing devaluations, dumping, and retaliatory protectionism, rendered them unreliable as a means of protection and most were soon converted to quotas. By 1937 Switzerland was the only nation employing tariff quotas on a wide scale [Heuser: 77].

Michael Rom’s book, *The Role of Tariff Quotas in Commercial Policy*, written in the mid-1970s, but based on work from the 1960s, provides a bridge between the pre-GATT descriptive studies and more analytical, Post-GATT studies. Rom is an advocate of employing tariff quotas as a means of providing generalized preferences for developing countries and most of the book is directed towards this end. There is a brief, partial equilibrium, graphical analysis of the economics of tariff quotas, and chapters on the use of tariff quotas by the EEC and the US. The US chapter focuses on the Woolen and Worsted Fabrics tariff quota; the EEC chapter is primarily concerned with intra-EC coal and steel trade.

When the European Economic Community (EEC) implemented its Generalized System of Preferences (GSP) it granted tariff preferences to qualifying developing countries as tariff quotas. The tariff quotas allow imports up to a base volume at the GSP rate; for imports exceeding this volume the Most Favored Nation (MFN) rate is applied. The tariff quota provides a direct safeguard against a (preferential) import surge. [Murray (1977): Ch.4]

There is a limited analytical literature on the economics of tariff quotas. Anderson (1989) and Young and Anderson (1980) explore some aspects of tariff quotas in the context of the non-equivalence of tariffs and quota literature in the late 1970s early 1980s.

Auctions of quota rights with and without imperfect competition

Bergsten et al. (1987) makes a strong case for auctioning quota. The book framed the issue in the hope of influencing the agenda for the Uruguay Round. It provides an excellent survey of how auctions have been implemented in trade and non-trade contexts. Most of the text remains timely and is a good introduction for a non-specialist.
Section 4.2 above discusses the literature on auctions, particularly the work of financial economists. International trade economists have been attracted to quota auctions as a means to examine (sometimes even empirically) imperfect competition. Most of the imperfect competition and international trade literature assumes a monopolistic or imperfect market as a starting point. The more interesting question is whether markets are imperfect and, if they are, what barriers to entry, transactions costs or natural advantages account for the persistence of incumbent rents. Krishna (1990, 1993) has established a program of research on the (in)efficiency of auctions under imperfect competition. The *Journal of International Economics* regularly publishes articles on the topic.

---

### Appendix D

#### TRQ Information on the Web

The WTO has a Document Dissemination Facility (ddf) where one may download country notifications to the WTO, in addition to other public WTO documents. The website is


Select search. Searching by symbol is probably the most efficient method, but one must first decipher WTO nomenclature. One can use the ‘symbols’ button browse the WTO code system. Most TRQ notifications can be found by searching for the following code:

G/AG/N/ member

Which, in natural language, reads: trade in Goods / Agriculture / Notifications / member code

A search for this code will return all country notifications on the topic. TRQ notifications are usually in the first or second notification. Often there are addenda to the initial notifications. The initial TRQ notifications for the United States are found in:

G/AG/N/USA/2, G/AG/N/USA/2/Add.1 and G/AG/N/USA/2/Add.2

Fill rates are also to be notified annually. For the United States, notification 3 reports TRQ fill for 1995 and notification 16 reports TRQ fill for 1996 and 1997.

Member country codes have three letters, but there are some false cognates. For example, CHE is Switzerland: CHL is Chile; and ISL is Iceland: ISR is Israel. Table 1 of this paper lists the codes for countries notifying TRQs.
Notifications of TRQ administration are reported in Table MA(1):

Table MA(1) MARKET ACCESS
Commodity description and tariff line(s)

(a) COUNTRY ALLOCATIONS: (describes method or reference to documentation)
(b) IMPORTER ALLOCATIONS: (describes method)
(c) OTHER ARRANGEMENTS: (description or reference to documentation)
(d) OTHER INFORMATION: (various)

Specific country allocations are often included in an annex to the table, generally part of the same document.

Notifications of quota fill (the volume imported in-quota) are reported in table MA(2):

Table MA:2 MARKET ACCESS

Description of products
Tariff item number(s) encompassed in product description
Tariff quota quantity for period in question Quantity
In-quota imports during period Quantity

The in-quota and over-quota tariffs \( \{t, T\} \) are notified to the WTO in a member’s UR tariff schedule.

Many UG tariff schedules are available on-line. For example, FAS - the Foreign Agricultural Service of USDA - has nice WTO tariff schedule search engine. Its website is:

http://www.fas.usda.gov/scriptsw/wtopdf/wtopdf Frm.idc

Much of the key TRQ information is buried in notes to the tariff schedule and these are often in a separate document.
References


MacDonald, James, M., Charles R. Handy, and Gerald Plato. 1998. *Food Procurement by USDA's Farm Service Agency*. Economic Research Service, AER No. 766.


Oskam, Arie, J. 1991. *Quota Mobility and Quota Values and Their Influences on Structural Development of Dairy Farming*. OECD.


