ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mönnich, Christina

Working Paper

Tariff rate quotas: Does administration matter?

Discussion Paper, No. 16

Provided in Cooperation with:
Justus Liebig University Giessen, Center for international Development and Environmental Research (ZEU)

Suggested Citation: Mönnich, Christina (2003) : Tariff rate quotas: Does administration matter?, Discussion Paper, No. 16, Justus-Liebig-Universität Gießen, Zentrum für Internationale Entwicklungs- und Umweltforschung (ZEU), Giessen

This Version is available at:
https://hdl.handle.net/10419/50488

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Tariff Rate Quotas:
Does Administration Matter?

by

Christina Mönnich*

No. 16, Giessen, December 2003

* Dipl.-oec. Christina Mönnich, Center for International Development and Environmental Research (ZEU), Justus-Liebig-University Giessen, Otto-Behaghel-Str. 10 D, D-35394 Giessen.
Christina.Moennich@zeu.uni-giessen.de,
http://www.uni-giessen.de/~gm1068.

Financial support by the VolkswagenStiftung is gratefully acknowledged.

CONTENTS

1 INTRODUCTION 1

2 THEORY OF TRQS AND THEIR ADMINISTRATION 2

3 RESULTS 7

4 THE MODEL 11

5 CALCULATION OF PARTIAL EFFECTS 15

6 CONCLUSIONS 16

7 REFERENCES 17

Tariff Rate Quotas: Does Administration Matter?

At the Uruguay Round, tariff rate quotas (TRQs) were intended to serve two purposes: first, to prevent that tariffication would lead – at least on the short term – to a deterioration of market access and second, to create new, minimum market access. The since then observed fill rates do not match with these intentions, being often rather low. A wide-held suspicion explains this with tariff quota administration imposing an extra barrier to trade. It is the aim of this paper to test whether different administrative methods do indeed contribute to explain variation in fill rates and if so, how. A censored regression model for panel data was developed and applied to the EU's TRQs for the years 1995–2000. The data supported the presumption that administration matters, sometimes in surprising ways.

1 Introduction

Since they were institutionalized in the Uruguay Round, tariff rate quotas (TRQs) have become an important market access instrument in agricultural trade. All OECD countries, except Turkey, have scheduled TRQs at the WTO (OECD 2002, p.20). These often account for a substantial share of tariff lines (OECD 2002, p.34). However, complaints about this trade instrument abound and consequently, TRQs are a prominent issue in the ongoing negotiations on agricultural trade liberalization. In particular, TRQ administration is often held responsible for the many low or medium fill rates that can be observed in most countries (WTO 2001). Several authors have raised this point, for instance Skully (1999), IATRC (2001), Boughner & de Gorter (1999). In the WTO talks, proposals for reform have been submitted by the US (WTO 2000*a*) or different developing countries –see Matthews & Laroche Dupraz (2001) for an overview. The Harbinson proposal also picked up the issue (WTO 2003).

This paper aims to contribute to the TRQ debate by presenting the results of an econometric model that was developed to test, first, which factors contribute to explain variation in fill rates and whether TRQ administration belongs to them. Second, if different methods of TRQ administration have an impact on fill rates –as, in fact, they do– the empirical analysis should indicate whether it is comparatively positive or negative. By applying the model to the EU TRQs for the period 1995–2000, these questions could be answered at least for this important agricultural import market, where TRQs play a substantial role. So far, there has been very little empirical work on TRQs in general and none on TRQ administration in particular. The OECD (2002) has modelled TRQs in a modified version of its Aglink model, but has abstracted from modelling the different methods of TRQ administration. Other contributions evaluated these administrative methods based on theoretical considerations or models (see Skully (1999), de Gorter, Falk & Hranaiova (2003), Chau, de Gorter & Hranaiova (2003) or Gervais & Surprenant (2000)) which allowed only few definite conclusions.

In order to apply an econometric model to TRQ data, some conceptual problems had to be overcome. The central question was how to combine a method suitable for censored regression with the panel structure of the data. The latter had to be exploited in order to avoid the inconsistency caused by an important omitted variable. The solution found is simple and appears to work very well. Thus, it could be a method of interest to many other applications with a similar data structure.

In the next section, the main characteristics of TRQs and their administration will be shortly presented. For a more extensive treatment, see, for instance, Skully (1999), OECD (2002), Mönnich (2003) or IATRC (2001). The emphasis of section 3 is on the main results of the model, while methodological issues are postponed to section 4. The presentation of partial effects in section 5 is a mixture of methodology and results. Section 6 concludes the paper.

2 Theory of TRQs and Their Administration

Tariff rate quotas set a quantitative limit on the availability of a tariff reduction. That is, a comparatively low in-quota tariff t_{in} is levied only up to the quota quantity QQ. The out-of-quota tariff t_{out}, in contrast, is higher, sometimes much higher than t_{in}, but its applicability is not limited (Figure 1). The out-of-quota tariff is the regular, most favored nation tariff bound in the schedules of WTO members. Many of these tariffs were bound for the first time in the Uruguay Round and were a result of the so-called tariffication, the conversion of non-tariff barriers into tariffs. Since quotas and other non-tariff barriers were far more common in agriculture than in other sectors, this was considered a significant success of the Uruguay Round Agreement on Agriculture (URAA).

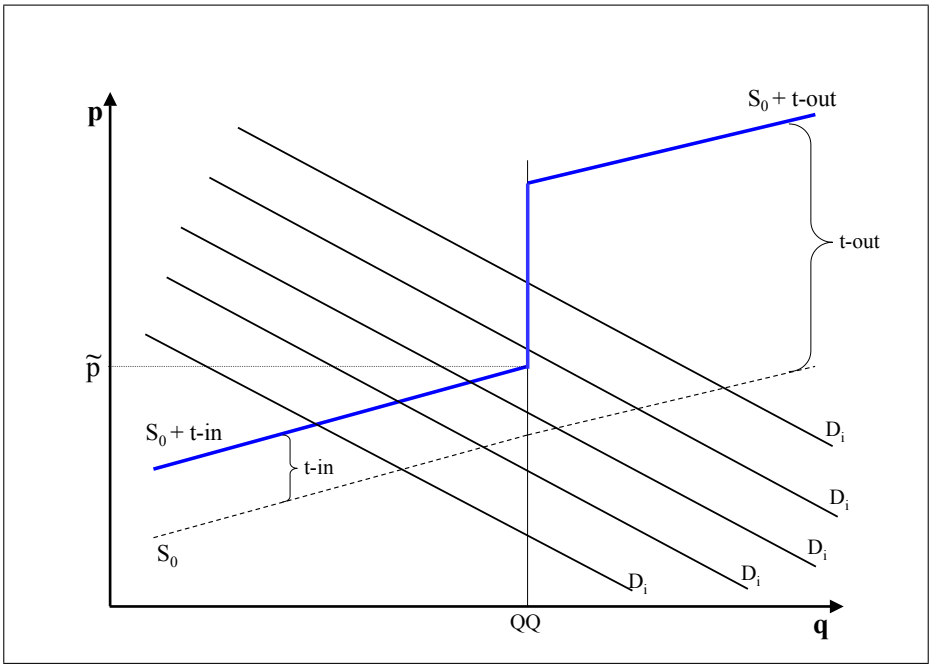


Figure 1: TRQ with Varying Demand (Large Country)

However, it became soon apparent that tariffication was often "dirty", because tariffs yielded a higher level of protection than had prevailed before (Ingco 1995). Or, said in more technical terms, the resulting tariffs were higher than the quota-equivalent tariff.[1] Therefore, so-called current access TRQs were intended to prevent that the

[1] Even if conversion had always yielded the quota-equivalent tariff, the latter would still not have reflected the pre-Uruguay-Round level of protection very well. This is due to the reference period of 1986–88 laid down in the Modalities for tariff conversion, which were years of exceptionally low world market prices and, accordingly, exceptionally high levels of effective protection in countries applying quotas (Boughner & de Gorter 1999).

trade round would actually lead to a decrease in trade. By applying lower tariffs to previously traded quantities, TRQs could help ensuring that at least the status quo was maintained. And since obligations to reduce the bound, out-of-quota tariffs were also part of the Agreement on Agriculture, future trade expansion was in prospect. The URAA further stipulated that market access equivalent to three percent of domestic consumption should be provided, if trade had not reached this level before. In order to comply with this obligation, many countries applied TRQs, so-called minimum access TRQs.

In sum, TRQs were institutionalized by the URAA with the aim to serve two purposes: First, current access TRQs were a safeguard of tariffication, which often resulted to be quite dirty due to the resistance to agricultural trade liberalization in many countries and the fact that the conversion of non-tariff barriers is, in practice, no exact science. So TRQs should secure current market access. Second, minimum market access should be accorded whenever trade had not reached the mark of three percent of domestic consumption prior to the negotiations. Here, TRQs provided an instrument to implement the obligation to liberalize while safeguarding, to a good degree, against import surges. In short, TRQs were meant to support the liberalization of formerly very protected markets.

Given this origin, it could be expected that the quotas of the TRQs would easily fill. However, this was often not observed in the years after the round. Instead, a rather mixed picture evolved with average fill rates that did not meet the above expectations. Three causes for quota underfill are most probable. First, there can simply be low demand. If demand is not just temporarily, but consistently low, it can be suspected that the product definition is too narrow or an unattractive market has been deliberately chosen. Second, the in-quota tariff t_{in} can be too high, rendering imports too expensive. Third, quota administration can impose obstacles to importation or cause higher transaction costs.

Quota Administration

In the academic literature and the political debate, the term quota administration comprises a whole range of quite different measures. There are first of all those administrative measures that are a constituent part of a TRQ, in the sense that their aim is to solve the rationing problem inherent in a quantitative restriction. They regulate importers' access to the quota. In contrast, country-specific allocations govern quota access on the supply side. Though an import quota can be effective without them, they are often applied in practice. Finally, there are often regulations in force that inhibit trade, or change its distribution, but which do not actually belong to quota administration: they can be applied irrespective of whether a quota is in force or not. Examples are discriminating practices against foreign import firms or requirements on how raw products are to be processed.

Here, the focus of attention will lie on transaction costs and x-inefficiency caused by rationing. As Skully (1999, p.4) puts it, on the point: "Rationing is the essence of TRQ administration". The reason is that the quota inhibits the market mechanism to balance supply and demand for any price higher than $\tilde{p}$, where importers are willing to purchase a larger quantity than QQ and exporters want to sell a larger quantity than QQ. Therefore, a rationing system must be established that determines who is to import under a quota. As already mentioned, TRQ administration often also regulates which exporting countries may supply under the quota, even though it is sufficient that the quantitative limit is imposed on just one market side.

Of course, if the quota does not fill, it is not necessary to impose its quantitative limit. But since this is not –or should not, at least– be known ex ante, the rationing system must be in place anyhow. This is also the case if the quota is not effective – or binding, in technical vocabulary– any more, that is if out-of-quota imports occur. Then, importers can usually be expected to rather pay the lower in-quota tariff than the higher out-of-quota tariff, and the rationing system must determine who of them will be the lucky importers. If, however, the spread between the two tariffs is not very large, but transaction costs caused by quota administration are, all importers are rather unlucky.

Transaction costs specific to quota administration are, for instance: costs of getting informed about the administrative method applied and the requirements to be met in order to qualify for quota access; costs due to the paperwork and time necessary for the application process; costs for hedging – as far as possible – against the insecurity of not getting quota access for the firm's profit-maximizing quantity; costs of speeding up the importation if the order of application is relevant or the end of the quota period is near; costs of finding business partners in a given export country, if importers are not free in their sourcing decisions; opportunity costs of the securities that often have to be lodged for the case that import licenses are not used.

Besides such transaction costs borne by private business, costs of the state bureaucracy borne by the taxpayer are also due to quota administration. However, the first are of greater interest here, because they have a possible impact on trade flows: if different methods of quota allocation cause transaction costs of differing magnitudes, then some methods will be preferable to others – unless this advantage is eliminated by higher x-inefficiency.

So-called x-inefficiency means that firms produce or supply services at higher than optimal costs. This is usually associated with imperfect competition. Besides the fact that rationing has an impact on competitive conditions, a binding quota drives a wedge between marginal prices of buyers and sellers. The consequence is that trading firms can afford to be inefficient and real life rationing systems do not prevent this: allocation does not occur according to relative cost efficiency, but according to some other criterion.

In the following, the most frequently applied methods of quota administration will be presented briefly.

License on Demand (LD)

Importers request licenses for the quantity they plan to import. If total requests exceed the available quota quantity,[1] a pro-rata reduction is applied. If such a reduction is anticipated, importers have an incentive to exaggerate their requests. The EU requires security deposits for the license quantities allocated to each importer, so that these run a risk of loosing the deposit if their license quantity exceeds their import capacity. Thus, a certain disincentive to exaggerate requests is in force.

This method of import allocation does not select importers at all, it simply adjusts requested import quantities to the quota if necessary. Consequently, if the quota is binding and a price premium can be attained on the domestic market, i.e. if rents arise, it is possible that these are dissipated in inefficient firms not precluded from the market. But firms, irrespective of their relative competitiveness, have higher costs due to the high degree of uncertainty that this method of quota allocation entails.

First-Come, First-Served (FC)

This method is self-explaining. It puts a premium on time, because importers have incurred already a very high percentage of total expenses at the time of being late. In such a case, they have the choice between paying the higher out-of-quota tariff, store or dispose the good or, maybe, deviate it to another market. Either way, it is quite likely that they will suffer losses. If importers do not want to carry the risk of being too late, quota underfill can be the consequence. In this setting, the informational capacities of the import authorities become an important factor.

Importers will adjust by making provisions to increase their chances of timely quota access, for instance by investing in faster modes of transportation or by storing the commodity at the border before the quota period starts. Evidently, these are again costs caused by quota administration. Whether x-inefficiency adds to these costs depends, as Skully (1999) puts it, on the correlation between relative cost efficiency and place in the queue. Arguments for expecting it to be positive or negative can both be found.

Another disadvantage of this method of import-side rationing is that it can bias the distribution of exports. If there are no refinements to the procedure that take into account the distance from the origin of imports, nearby exporting countries are evidently favored.

[1] In the EU practice, licenses are usually allocated quarterly so that the quota quantity is –more or less– a quarter of the total yearly quantity.

Historical Importers (HI)

The allocation of quota access follows the distribution of trade of a previous reference period. For instance, importers' quota shares are determined by their respective average market shares of the preceding three years. This will obviously lead to a quite static distribution of trade. Newcomer shares can introduce some dynamics and are the only chance for other firms to enter the market where, otherwise, competition has been eliminated. The problem is to find the right size of the newcomer share: if it is too small, it will remain ineffective and prospective market entrants are discouraged from building up capacities. If its too large, historical firms will found subsidiaries that sign as newcomers. Since there will be regulations to prevent this, resources will be used to enforce them and to circumvent them anyhow.

Besides these problematic features, this method of quota allocation has a clear advantage, at least for those importing firms favored by the system, i.e. by those who were already in the market before imposition of the TRQ or quota: it makes planning quite easy. Both compared to free trade and to other methods of quota allocation, calculation is much facilitated.

If high fill rates are a priority to regulators, this method of quota allocation is a good choice. Even in years where imports are not very profitable, importers will think twice before reducing their imports, because this reduces their future reference quantities and therefore quota shares.

Other Methods of Quota Administration

Other methods of quota administration like auctioning, state trading entities or producer groups are not discussed here. They are seldom applied globally and not at all by the EU. Interested readers may refer to Skully (1999) or Mönnich (2003). In a report on TRQ administration compiled by the WTO committee on agriculture (WTO 2000*b*), applied tariffs are also categorized as a method of quota administration. But if the in-quota tariff is simply applied on all imports, there is no effective quota in force and therefore, the term is not very accurate. However, almost half of all TRQs notified at the WTO in 1999 were applied tariffs (WTO 2000*b*, p.5).[2]

[2]This can be explained as safeguard measures. Once a TRQ is scheduled in accordance with trade negotiations, a country has the option to change the method of quota administration later and thus make the quota effective later. The incentive to do so arises from the general rule that WTO member countries have to compensate adversely affected trade partners if they apply more protectionist policies than those they have scheduled. In this case, the policy change is not more protectionist only from a legalist point of view.

Country-Specific Allocations

Country-specific allocations are quota shares that the importing country reserves for specific exporting countries. They are therefore a discretionary distribution of access to the quota that the import country imposes on the supply-side. Usually, at the time of imposition of the quota, this distribution is based on historical trade flows and not changed very frequently afterwards.[3] Therefore, the arguments are very similar to those for (or against) the method of historical importers, even though the allocation of market access occurs at a more aggregated level. Most likely, political considerations will play a major role for the decision whether country specific allocations should be imposed or changed. The need to compensate trade partners for the introduction of the quota or the wish to pursue developmental goals will often dominate over concerns about economic efficiency.

Since country-specific allocations limit importers in their sourcing decisions, the bargaining position of exporters is improved. Export certificates even magnify this effect. Then, importers do not only need an import license for the quantity to be imported, but also a matching export certificate. This requirement counterbalances the bias that otherwise disadvantages exporters because they can enter into negotiations only with importers that hold import licenses.[4] From all countries imposing TRQs, only the EU makes extensive use of export certificates.[5] In fact, they are highly correlated with the incidence of country-specific allocations. So it can be expected that they have a negative impact on fill rates, because they will cause higher costs to importers and because they capture the effect of country-specific allocations. These, for the reasons mentioned above, are also not expected to increase efficient allocation of market access.

3 Results

The discussion so far can be summarized as follows: The fill rate of a TRQ is determined by its quota size and its tariffs relative to import demand. Furthermore, quota administration will hardly be neutral on the operation of TRQs. This section focuses on the results of the model that was developed to test these hypotheses and to find out more about individual effects. Only the most important characteristics of the model will be described here, while a more formal methodological discussion will follow in section 4.

[3] For instance, the predecessor of a US-TRQ for sugar, a quota, was introduced in 1934. Major reallocation occurred, on average, every 15 years and were always prompted by an economic or political shock (Skully 1999, p.22).

[4] This argument does not apply if import side market access is administered according to first-come, first-served. But, for the EU, this is the case only twice.

[5] Besides the EU, the US requires them twice, Canada just once (WTO 2000*b*).

The dependent variable is a vector of fill rates as notified by the EU at the WTO. Since only fill rates up to 100% are reported, the regression is censored. Often, however, the import statistic reports higher import quantities. The excess quantities are assumed to be "parallel" imports and not out-of-quota imports for which the out-of-quota tariff has been paid. This assumption is based on the observation that out-of-quota tariffs are usually substantially higher than the corresponding in-quota tariffs – despite the reductions proscribed in the URAA (see Appendix). Therefore, out-of-quota imports are assumed to be negligible so that the out-of-quota tariffs do not enter the model.

The assumption of negligible out-of-quota imports is in line with the assessment of the OECD (2002). That is, in the many instances where actual imports exceed the quota size, the exceeding quantities are believed to occur under preferential regimes of which the EU has plenty, like the Europe-Agreements or the Euro-Mediterranean Agreements.

Since such imports can be observed frequently and on a large scale, they enter the model through the variable qtom. This relates the quota quantity QQ to all imports of the corresponding tariff lines. Its effect will be discussed further down.

Another variable that warrants an explanation is ν. It is a between-group effect that captures heterogeneity between the different TRQs. To a large degree, this heterogeneity is expected to be caused by an important missing variable, the relative quota size. The absolute quota size is of course known, but it is a meaningless quantity if it is not compared to the size of the domestic market. Given that most TRQs are defined on a quite differentiated product level, a large number of market studies would have been necessary to get even close to this measure.

Besides relative quota size, ν also captures product-specific peculiarities not explained elsewhere. Therefore, it helps avoiding the inconsistency that is expected if an important variable is missing. It also helps improving the precision of the estimator that is expected to be low if there is too much unexplained variability. It was possible to estimate ν thanks to the two-dimensional panel data structure of the data. In the cross-section dimension, the different TRQs for different products form the so-called groups. The period 1995–2000 for which observations were collected provides the time-series dimension. In the cross section, $N \leq 87$, because some TRQs were introduced later than 1995. Accordingly, $T \leq 6$.

Table 1 summarizes the results of the model, together with a short description of the variables.

The import price has a coefficient with a high level of significance, and the coefficient is positive. There are two arguments to explain this, based on the panel structure of the data. In the cross section dimension, it means that TRQs for higher-valued products tend to be better used. Given the relative high importance of transport costs for agricultural products, this is plausible. With respect to variation over time, the positive coefficient is only explicable with imports being demand-driven and the EU being the large country

Table 1: Main Model Results

variable	coef.	t	description
fr			dependent variable; notified imports according to WTO notifications over quota size according to EU regulations
ν	1.135	25.08	group effects, see section 4
mprice	.090	10.60	cif import prices according to the trade statistic
qtom	-.040	-6.55	quota size in relation to all imports of the corresponding CN-codes, according to the import statistic
fc	.178	4.32	dummy for first-come, first-served ; source: WTO (2000*b*)
hi	.172	8.47	dummy for historical importers; source: WTO (2000*b*)
ld			license on demand; reference administration method for fc and hi; source: WTO (2000*b*)
xcert	.107	4.31	dummy for export certificates; highly correlated with country-specific allocations; source: WTO (2000*b*)
cons	.715	35.19	constant
	pseudo R^2	.46	pseudo-$R^2 = 1 - \frac{\text{min. sum of dev.}}{\text{raw sum of dev.}}$

of the textbooks.[6] Figure 1 illustrates this. If, unlike presented there, supply was totally elastic, the import price would not vary with demand shocks. If supply instead of demand varied, the coefficient would be negative.

This result could be interpreted to imply that imports under TRQs are used to satisfy varying demand. Since total demand is usually rather stable, these variations in import demand will primarily stem from varying supply of domestic producers. In other words, imports under WTO-TRQs are primarily a complement to domestic supply. Nevertheless, the EU import market is so large that its supply function is positively sloped.

The next argument is closely related to these arguments. The negative coefficient of qtom means that the higher total imports into the EU, the higher the fill rates of the WTO-TRQs tend to be. As mentioned above, many of these TRQs are not the only possibility of entry to the EU market, but one of several. This motivated the inclusion of the variable qtom which measures QQ/(total imports). Its evaluation shows that the incidence of such parallel import regimes is not negligible: For 270 observations, qtom < 0.8, implying that total imports are at least 25 percent ($1/0.8 = 1.25$) larger

[6] The two-step procedure to account for the group effects (section 4) distills the cross-section dimension. The coefficient of mprice is positive in both regressions, the between-groups and the pooled one. This indicates that the two separate arguments can both be valid.

than the quota quantity. The mean of these observations is 0.3. Since it is impossible to link import quantities to the regime under which they took place, it cannot be excluded that some of these are "real" out-of-quota imports, for which the out-of-quota tariff has been paid. As has been already argued, this is not expected to be the case very often, given the generally high level of out-of-quota tariffs. In any case, a high value of qtom signals an attractive market.

Conversely, if qtom < 0.8, the WTO-TRQ is probably the only opportunity to enter the EU market. Since the mean of these observations is 8.2, which translates to an average fill rate of 0.12, this market access opportunity is generally of low value to market participants. In general, the negative coefficient of qtom can be plausibly interpreted to mean that an attractive market tends to lead to higher fill rates of the WTO-TRQs, though these might be small compared to other import opportunities.

Coming to the administrative variables, they all have significant and positive coefficients. The dummies for historical importers and first-come, first served compare to license on demand, while the reference situation of export certificates is that none are required. This essentially means that there are no country-specific allocations, because these two variables are highly correlated. It was expected that the higher bureaucracy caused by export certificates would lead to lower fill rates. Likewise, it was expected that the reduced competition caused by country-specific allocations would lead to lower fill rates. However, neither of these expectations was supported by the data: the coefficient estimate of xcert is positive and not negative.

That reduced competition would cause lower fill rates could also be expected for historical importers. Here, too, the results of the model do not confirm this expectation. In both cases, the guaranteed and calculable quota access reduces insecurity. This apparently dominates any negative effects that either country-specific allocations or the method of historical importers might have anyhow. That is: the positive coefficients that both dummies have do not imply that these methods are efficient. These methods of quota administration do not ensure that the lowest cost firms get quota access or that production is cost efficient in the first place. The positive coefficients only suggest that export certificates respective country-specific allocations lead to higher fill rates and that the method of historical importers leads to higher fill rates than the method of license on demand.

As far as the method of first-come, first-served is concerned, it was a priori not really clear whether it implies a lower or higher level of insecurity than license on demand. But apparently importers have adapted to the system and can better deal with it than with the randomness of license allocation of the latter. Therefore, the method of license on demand performs poorly compared to the other methods of quota allocation applied by the EU. The formal model in de Gorter et al. (2003) provides a possible explanation: this method restrains the most efficient firm first, while less competitive firms can get closer or even reach their optimal import quantity. Adding imperfect information and the insecurity it entails, it is not surprising that fill rates are comparatively low.

Taking everything together, administration matters. It matters besides import prices, parallel imports and the group effect. This latter is supposed to capture mainly the unobservable effect of relative quota size, but it will also capture any otherwise unexplained heterogeneity of the products in question.

In contrast to these findings, it is remarkable that the in-quota tariff does not appear to make a significant contribution to explain fill rates. Furthermore, it was surprising that the relative price difference between the domestic and the import price of the product in question did also not yield a coefficient with a satisfying level of significance. These results will be discussed in more detail further down, in connection with the development of the model. Therefore, the properties of the model should be described first.

4 The Model

When devising a model that could be used to test which variables determine the fill rates of TRQs and whether quota administration belongs to them, three things had to be taken into account:

1. A **variable** that is expected to be a very important determinant for both fill rates and prices, namely the relative quota size, is **missing**. More precisely, the relative quota size is de facto unobservable, because data on the size of the domestic market are usually not available, given the very specific product definitions for most TRQs.

2. The data on fill rates are **censored**. This is naturally due to the very purpose of a quota, as can be seen in Figure 1. In the inelastic stretch of the effective supply curve, differing demand conditions do not prompt changes in import quantities. As has been discussed above, it is assumed that out-of-quota imports are, in general, not economically viable. Where imports exceed those notified under WTO-TRQs, the excess quantities are assumed to take place under other, bilateral import arrangements of the EU.

3. Possible **simultaneity** must be taken into account. If prices are expected to affect fill rates, fill rates might as well affect prices. This appears even more probable given that the discussion from above has shown that quota administration can be expected, in many instances, to affect both prices and fill rates.

Omitting such an important **variable** as relative quota size would have risked inconsistent and unprecise estimators. Therefore, as already mentioned, the idea was to take advantage of the panel structure of the data and introducing the group variable ν:

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + \nu_i + \varepsilon_{it}, \qquad \varepsilon_{it} \mid \mathbf{x}_i, \nu_i \sim N[0, \sigma_\varepsilon^2] \tag{1}$$

The classical panel techniques of fixed effects and random effects differ in the method of estimating the ν_i and in the assumptions concerning the correlation between the group

effect and the regressors (Wooldridge 2002, pp. 251–52). In a random effects model, $\boldsymbol{\nu}$ is assumed to be independent of the regressors and it is assumed to be normally distributed, as $\nu_i \mid \mathbf{x}_i, \sim N[0, \sigma_\nu^2]$. The fixed effect model treats the ν_i as group-specific constants.

Panel regression had to be combined be with **censored regression**, because regular regression techniques that do not account for data censoring yield, again, inconsistent estimators. The standard approach is the tobit model which is estimated with maximum likelihood. The unobserved effects tobit model following Wooldridge (2002, pp.540–42) combines a tobit with a random effects model. Its loglikelihood function is maximized with respect to $\boldsymbol{\beta}$, σ_ν and σ_ε, maximizing the joint probability of all groups. This model is already implemented in Stata7, which approximates the function and calculates the derivatives with a M-point Gauss-Hermite quadrature (Stata Press 2001*a*, p.28). For this method to work, however, $\rho = \sigma_\nu^2/(\sigma_\varepsilon^2 + \sigma_\nu^2)$ must not be too large. Otherwise, the function is not well-approximated by a polynomial (Stata Press 2001*b*, pp.477–478). Here, the estimated $\rho \approx .73$ is to be blamed for the numerical instability that was confirmed by a sensitivity test for numerical soundness provided by Stata (Stata Press 2001*a*, pp.28–32).

Honoré (1992) proposes a fixed effects model for censored regression settings. However, fixed effects rely, for estimation, on the so-called within-group variation, that is, on variation within a cross-section unit over time. Since the method of quota administration, once fixed, is usually not changed for given TRQs,[7] this approach was also not practicable.

Following approach to solve the omitted variable problem turned out to be practicable: The model was estimated twice. In the first step, a pure cross section regression was performed with the averages per TRQ over time. That is, all variation over time was eliminated, so that estimation was based solely on between-group variation. The residuals from this regression were used, in the second step, as proxy for the group effects ν_i. Though they do not only capture the effect of the omitted variable "relative quota size", but any product peculiarities as well, they served their intended purpose quite well. This was first, to avoid the inconsistency of a regression with an important regressor missing and second, to get a better estimate for σ^2, because an inflated $\hat{\sigma}^2$ (that incorporates $\boldsymbol{\nu}$) makes confidence intervals shrink. As can be seen in Table 2, the second effect turned out to be more important. It compares the estimates of the first, in-between regression with the pooled regression including the proxies for ν.

Since both regressions use pure cross-section techniques of estimation, the censored regression model does not need to be, simultaneously, a panel data model any more. But instead of using the regular cross-section tobit approach, a median regression adapted to

[7] In fact, the banana TRQ is the only exception. Here, some changes were prompted by different WTO rulings (see Herrmann, Kramb & Mönnich (2003)). However, these changes occurred within a given method of administration, so that the dummies do not change.

Table 2: The Effect of the ν-Correction

Comparison of Coefficients			**Comparison of Standard Errors**		
	in-between (n=79)	pooled (n=428)		in-between (n=79)	pooled (n=428)
β_1	.071	.090	se_{β_1}	.04	.008
β_2	-.051	-.040	se_{β_2}	.012	.006
β_3	.178	.172	se_{β_3}	.085	.020
β_4	.114	.178	se_{β_4}	.13	.041
β_5	.082	.107	se_{β_5}	.090	.025

data censoring following Powell (1984) was applied: unlike the tobit, it does not hinge on distributional assumptions and is robust to heteroscedasticity.

As usual, the model formulation is based on an uncensored latent variable y^*. Censoring occurs from above, at fill rates of 100 percent.

$$\begin{aligned} y_{it}^* &= \boldsymbol{\beta}'\mathbf{x}_{it} + \varepsilon_{it}, \\ y_{it} &= y_{it}^* && \text{if } y_{it}^* < 1, \\ y_{it} &= 1 && \text{if } y_{it}^* \geq 1, \end{aligned}$$

The central idea is based on the aim of median regression, which is to estimate the expected median, conditional on the regressors: Assuming that the error terms have zero median, the probability that $y_i = 1$, which is equivalent to the probability that

$$y_{it}^* = \boldsymbol{\beta}'\mathbf{x}_{it} + \varepsilon_{it} \geq 1,$$

is less than one-half if $\boldsymbol{\beta}'\mathbf{x}_{it} < 1$. Likewise, if $\boldsymbol{\beta}'\mathbf{x}_{it} > 1$, the probability that

$$y_{it}^* = \boldsymbol{\beta}'\mathbf{x}_{it} + \varepsilon_i \geq 1$$

exceeds one-half. Then, the probability that $y_{it}^* = 1$ is equally greater than .5 and one will be the median of y_{it}.

Since the conditional median is the line through the data which minimizes the sum of absolute residuals, the objective function to be minimized with respect to $\boldsymbol{\beta}'$ is

$$1/N \sum_{i=1}^{N} |y_{it} - \min(\boldsymbol{\beta}'\mathbf{x}_{it}, 1)|. \tag{2}$$

If $\boldsymbol{\beta}'\mathbf{x}_{it} > 1$, the value of the term $|y_{it} - 1|$ is unrelated to $\boldsymbol{\beta}'$. Thus, minimization is performed only for observations for which $\boldsymbol{\beta}'\mathbf{x}_{it} < 1$. The regularity conditions that must be met for this procedure to work and for this censored least absolute deviations (CLAD) estimator to be asymptotically normal are quite general and not very restrictive (for more details, see Powell (1984) or Mönnich (2003)). In contrast to Powell's

Table 3: Results of the Hausman test

var	$\boldsymbol{\beta_n}$	$\boldsymbol{\beta_x}$	$\boldsymbol{\beta_n} - \boldsymbol{\beta_x}$	sqrt(diag($V_n - V_x$))
β_1(hat)	.088	.090	-.002	.006
β_2	-.04	-.040	-.0002	·
β_3	.177	.172	.005	009
β_4	.189	.178	.011	.022
β_5	.102	.107	-.004	.013
	χ^2_5	0.87	$P > \chi^2$	0.97

Note: The subscripts n and x stand for endogenous and exogenous, respectively.

suggestions, the variance-covariance matrix is bootstrapped. For more detail on how this and the above minimization task has been implemented using Stata7, see Mönnich (2003).

The issue of **simultaneity** turned out not to be a problem. A two-stage procedure in analogy to 2SLS was applied in order to eliminate possible correlation between the error terms and the regressors. For this purpose, the price variable was regressed on all exogenous variables. The fitted values were used to replace, in the second stage, the original observations of the price variable. The coefficient estimates of this second stage were very similar to the original ones and a Hausman test rejected the H_0-hypothesis that the vector of differences is significantly different from zero (see Table 3).

However, this should not be interpreted to imply that the price variable is in fact endogenous. The Hausman test simply indicates that there is no strong correlation between the error terms and this is what matters for consistent estimation. In contrast, the economic interpretation of various model results convincingly argues for simultaneity. In particular, most arguments why TRQ administration should matter apply equally to prices as to quantities. This strongly suggests that prices and quantities are determined simultaneously. In fact, two first stage regressions yielded good results: with the import price as regressand and with the price difference between the domestic and the import price (pdiff) as regressand.[8]

It was a particular surprise that the first-stage regression with pdiff as regressand yielded good results, while pdiff was not significant as a regressor for the fill rates, whether assumed to be exogenous or endogenous. Equally surprising was that the in-quota tariff appears to have a significant –and positive– coefficient only when pdiff was regressed on it. How does this all fit together? That neither the in-quota tariff nor pdiff have explanatory power for the TRQ fill rates could be interpreted to indicate that there is "water in the price transmission", meaning that domestic prices vary due to factors that have nothing to do with the import market. This, in turn, is consistent with the model results of the OECD (2002). According to those, an expansion of all quota quantities by 50% would not lead, in general, to a substantial decrease of domestic prices. Domestic

[8]It was not attempted to model prices in a two-stage procedure for reasons of identifiability.

policies are to be blamed for this: they are still in place for many products governed by TRQs, shielding domestic prices from world prices.

Consequently, it is plausible that the price difference between domestic and import prices does not –measurably– contribute to explain fill rates. These findings fit in well with the positive coefficient of the import price. This has been interpreted to imply that imports are mostly used to complement varying domestic supply.

So far, the focus of attention was on the question which of the possible impacts on fill rates were or were not statistically significant, how this could be explained and how this could be modelled in the first place. The magnitude of the effects was excluded from the discussion. In the next section, a tentative approach to obtain partial effects is presented.

5 Calculation of Partial Effects

In censored regression settings, the values of the coefficient estimates are not equal to the marginal partial effects. An adjustment is necessary that accounts for the change in probability that the observation no longer lies in the uncensored region. It follows that distributional assumptions are needed if this adjustment is to be made.

However, the CLAD regression achieves coefficient estimates without making distributional assumptions, and this was, besides its robustness, what made it attractive. But since it was unsatisfactory not to not to be able to get a better feel of the magnitudes of the effects, the model was re-estimated with tobit, making use of the already estimated group effects to account for the missing variable problem. This way, the difficulties of the random effects tobit mentioned above were avoided. This combination of methods seems not only justified by the interest on the approximate magnitude of the effects, but also by its results, as documented in Table 4. They indicate, contrary to ex-ante doubts in this respect, that the real distribution of the fill rates cannot be that far away from the normal distribution. Therefore, partial effects were tentatively calculated under the assumption of normality, with the formula:

$$\frac{\partial \mathrm{E}[y \mid \mathbf{x}]}{\partial x_j} = \Phi\left(\frac{1 - \boldsymbol{\beta}'\mathbf{x}}{\sigma}\right) \cdot \beta_j.$$

For evaluation, the means were used in $\mathbf{x}$. In $\boldsymbol{\beta}'$, the mostly lower CLAD estimates were used, based on the premise that it is preferable to underestimate an effect rather than to overestimate it. Thus, the adjustment factor is:

$$\Phi\left(\frac{1 - .804}{.256}\right) = .778.$$

Only for the dummy variables for historical importers and first-come, first served are the differences of the coefficient estimates appreciable. Most likely, the binary format

Table 4: Comparing the Coefficients of the CLAD and Tobit estimation

	CLAD		**tobit**	
var	coef.	t	coef.	t
nu	1.135	25.08	1.269	21.53
mprice	.090	10.60	.092	9.22
qtom	-.040	-6.55	-.043	-12.27
hi	.172	8.47	.358	5.93
fc	.178	4.32	.261	6.55
xcert	.107	4.31	.092	2.89
cons	.715	35.19	.763	27.82
	Pseudo-R^2	.55	Pseudo-R^2	.83
			σ	.256

The value of the $\chi^2_6 -$ statistic of the LR-test for the tobit model is 404.08 with Prob $> \chi^2 =$ 0.00.

of the variables is responsible for these differences. For binary variables, relative effects were calculated according to the formula

$$g_{raw} = \frac{\mathrm{E}[y \mid x_j = 1] - \mathrm{E}[y \mid x_j = 0]}{\mathrm{E}[y \mid x_j = 0]}.$$

and then adjusted for data censoring. The results of all these calculations can be seen in table 5. At the mean of the data, prices matter, and administration matters.

Table 5: Estimated Partial Effects

var	mean	coef.	$\frac{\partial \mathrm{E}[y\|\mathbf{x}]}{\partial x_j}$	δ	g
nu	-.09	1.135			
mprice	1.88	.09	.07	.16	
qtom	1.51	-.04	-.03	-.06	
hi	.08	.172			.17
fc	.22	.178			.18
xcert	.27	.107			.11
cons		.715			

6 Conclusions

It was the aim of this work to develop an econometric model that explains the fill rates of the EU's TRQs fairly well, at least well enough to be able to test whether

quota administration plays a role as well. The discussion of the rationing problem that is immanent in any quantitative restriction led to the quite general expectation that somehow, it would. However, a closer theoretical analysis of the different methods of quota allocation did often not allow a conclusive argument whether the effect is expected to be positive or negative. The results generally confirm that quota administration is indeed not neutral, but sometimes they do not confirm common presumptions. In particular, allocating quota shares according to historical trade flows, be it on the import side for firms or on the export side for countries, does not appear to cause quota underfill. The inefficiencies attributed to this way of rationing cannot be as serious as expected, at least as far as the WTO-TRQs of the EU are concerned.

However, it should not be concluded that TRQs would generally perform better if their quota administration was changed to the methods that seem to perform relatively better. First, such changes will probably be disruptive, because market participants have adapted to the method of administration in place. The example of the banana TRQ strongly suggests so (Herrmann et al. 2003). More importantly, the familiar critique against TRQs is still of great importance. That is, out-of-quota tariffs are often prohibitive with the consequence that TRQs are really quotas; quota rents are dysfunctional to competitive markets; the rationing problem can never be solved satisfactorily and the bureaucracy needed to administer quotas and to be administered is costly. For these reasons, the proclaimed purpose of TRQs to ease the transition from a quota to a tariff system should be taken seriously. Continuing to reduce out-of-quota tariffs would surely be very appropriate. Increasing quota size as well, especially of those TRQs that are consistently filled to 100 percent, improves market access immediately and has the potential of reducing the distortive effects of quota administration.

But whenever new TRQs are introduced with the justification of creating new market access, whether minimum or not, then these results should be taken into consideration. At the same time, analysis of the variable qtom has indicated that some TRQs are of little commercial interest. Restrictive product definitions often seem to be responsible for this. Therefore, provisions to prevent such definitional restrictions still need to be improved as well.

In sum, quota administration matters. So do the "classical" parameters of TRQs.

References

Boughner, D. S. & de Gorter, H. (1999), 'The Economics of 2-tier Tariff-Rate Import Quotas and the Agreement on Agriculture in the WTO'. Revised version of a paper presented at the Annual IATRC Meeting in December 1998.

Chau, N., de Gorter, H. & Hranaiova, J. (2003), '"Rent Dissipation" versus "Consumer Rent Appropriation with First-Come First-Served Import Quotas in Agriculture'.

Paper presented at the International Conference: Agricultural policy reform and the WTO: where are we heading? Capri (Italy), June 23-26, 2003.

de Gorter, H., Falk, J. & Hranaiova, J. (2003), 'The Economics of Administering Import Quotas with Licenses-on-Demand in Agriculture'. Paper presented at the International Conference: Agricultural policy reform and the WTO: where are we heading? Capri (Italy), June 23-26, 2003.

European Commission (1999), 'Commission regulation (ec) no 2204/1999 of 12 october 1999, amending annex i to council regulation (eec) no 2658/87 on the tariff and statistical nomenclature and on the common customs tariff', *Official Journal of the European Communities* .

Gervais, J.-P. & Surprenant, D. (2000), 'An Economic Investigation of the Import Licensing Methods and TRQs in Agriculture', *Canadian Journal of Agricultural Economics* **48**(4), 397–410.

Herrmann, R., Kramb, M. & Mönnich, C. (2003), 'The banana dispute: survey and lessons', *Quarterly Journal of International Agriculture* **42**(1), 21–47.

Honoré, B. E. (1992), 'Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects', *Econometrica* **60**(3), 533–565.

IATRC (2001), 'Issues in Reforming Tariff-Rate Import-Quotas in the Agreement on Agriculture in the WTO'. Commissioned Paper No.13.

Ingco, M. D. (1995), 'Agricultural Trade Liberalization in the Uruguay Round. One Step Forward, One Step Back?'. Policy Research Working Paper 1500, The World Bank.

Matthews, A. & Laroche Dupraz, C. (2001), 'Agricultural Tariff Rate Quotas As a Development Instrument', *Économie internationale* **87**, 89–106.

Mönnich, C. (2003), Tariff Rate Quotas and their Administration. Theory, Practice and an Econometric Model for the EU, PhD thesis, Justus-Liebig-Universitaet, Giessen, Germany.

OECD (2002), *Agriculture and Trade Liberalisation. Extending the Uruguay Round Agreement*, OECD Publications, Paris.

Powell, J. L. (1984), 'Least Absolute Deviations Estimation for the Censored Regression Model', *Journal of Econometrics* pp. 303–325.

Skully, D. W. (1999), 'The Economics of TRQ Administration'. Working Paper 99–6, International Agricultural Trade Research Consortium.

Stata Press (2001*a*), *Stata Reference Manual. Release 7*, College Station, Texas. Vol. 3.

Stata Press (2001*b*), *Stata Reference Manual. Release 7*, College Station, Texas. Vol. 4.

Wooldridge, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge, Massachussetts and London, England.

WTO (2000*a*), 'Proposal for Tariff Rate Quota Reform. Submission from the United States'. WTO Document symbol = G/AG/NG/W/58.

WTO (2000*b*), 'Tariff Quota Administration and Tariff Quota Fill'. WTO Document symbol = G/AG/NG/S/8.

WTO (2001), 'Tariff Quota Administration and Tariff Quota Fill. Background Paper by the Secretariat. Revision'. WTO Document symbol = G/AG/NG/S/8/Rev.1.

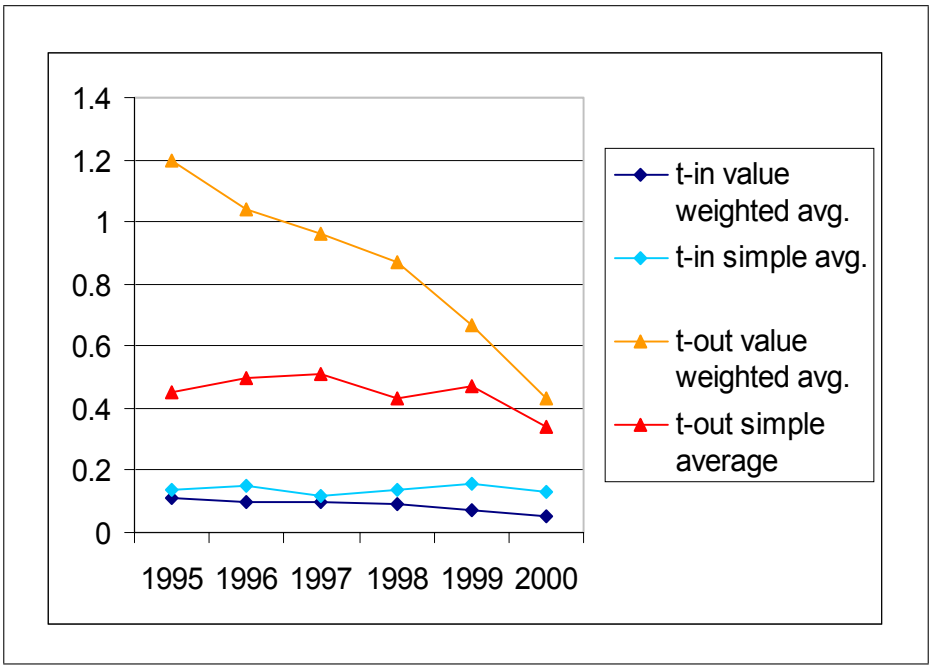
WTO (2003), 'Negotiations on Agriculture. First Draft of Modalities for the Further Commitments. Revision'. WTO Document symbol = TN/AG/W/1/Rev.1.

Appendix

Figure 2 illustrates that the tariff reductions negotiated at the Uruguay Round have shown results, while Figure 3 shows that individual out-of-quota tariffs are very often still multiples of the respective in-quota tariff.[9] Simple frequencies –i.e., each observation enters with equal weight– were compared with the distribution resulting from weighting each observation with its relative import value share. This comparison shows that the ratio t_{out}/t_{in} tends to be higher for TRQs of greater commercial importance.

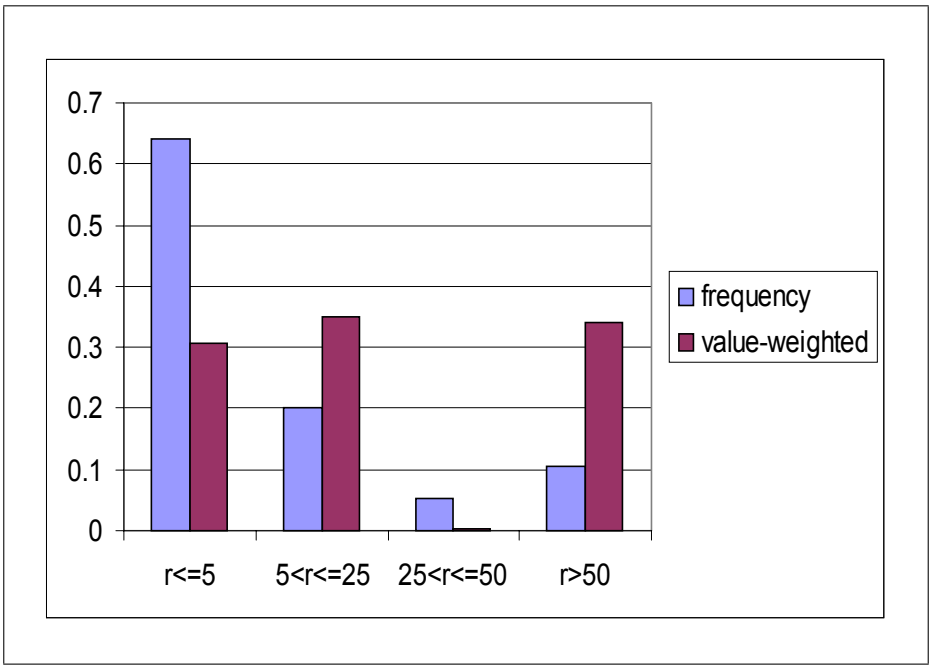
[9]Zero in-quota tariffs were replaced with a fictitious rate of one percent, in order to avoid losing 16 observations. Eleven observations are missing. These are mostly TRQs for fruits, for which no regular out-of-quota tariffs exist.

Figure 2: In-quota and Out-of-quota tariffs, 1995–2000



Sources: Own calculations based on various EC-Regulations, Eurostat.
Note: For TRQs with various tariff lines, import quantities were used for aggregation. Specific tariffs were converted to ad-valorem equivalents with average import prices. For the value-weighted average, each TRQ was weighted with its share of total import value (of all WTO-TRQs).

Figure 3: Ratios of t_{out}/t_{in}, frequencies and value-weighted distribution for 2000



Sources: Own calculations based on EC-Regulation (European Commission 1999), Eurostat.