ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master's Thesis Economics & Business: International Economics

Investigating the Impact of Sanitary and Phytosanitary measures on Non-Tariff Measure use

Anthony O'Dowd

429473ao

January 2020

Abstract

While there has been a general reduction in tariffs on trade since the creation of the World Trade Organisation (WTO) in 1995, there has been a rise in the use of Non-tariff Measures (NTM). Three major NTM types are Anti-Dumping Policy (ADP), Countervailing (CV) measures and Sanitary and Phytosanitary (SPS) measures. This paper investigates how the use and possible abuse of SPS measures influences other countries to change their use of NTM in retaliation. It aims to do this by answering the question: what are the responses to SPS measures by opposing countries, in the form of NTM (ADP and CV) use? This involves analysing every SPS measure, from the WTO list of SPS concerns – SPS measures which have been flagged as a potential violation of WTO rules - to investigate the change in NTM use each has caused i.e. the change in the number of NTM after, relative to before the SPS was implemented. This study finds that for 71.3% of SPS measures, there is no observable relationship between SPS and ADP (for CV it is 94.9%). But for the large tranche remaining for ADP measures (28.7%), SPS measures are shown to reduce other countries' use of ADP measures. Conformity Assessment (CA) SPS measures were found have a stronger effect than Product Characteristic (PC) SPS measures on reducing ADP use. The study also finds that groups of multiple SPS measures have the opposite effect, implementing an extra 1 SPS measure within a short time of the first, increases the response of ADP in return by between 0.409 and 0.925. For the remaining 5% of CV measures, no relationship was found with SPS.

1. Introduction

Under the influence of the World Trade Organisation (WTO) there has been a widespread reduction in trade barriers such as tariffs. However, Maskus and Wilson (2001) have noted there has been an increasing use of technical regulations by many countries, some of which may be using these as an alternative to more formal barriers to trade. These can be referred to as non-tariff measures (NTM). The United Nations (UN) defines NTM as a wide range of trade policy measures that influence international trade in goods by affecting quantities traded or prices (United Nations, 2020). The three major types of NTM are Anti-Dumping policy (ADP), Countervailing (CV) measures and Sanitary and Phytosanitary (SPS) measures. ADP are measures which countries can levy, under WTO rules, against imports which may threaten the domestic industry making the same product (Niels, 2000) . A country is said to be dumping when they export a product at a price lower than the price it normally charges in its home market. CV measures are also known as "anti-subsidy" measures. CV measures counter imports of products from firms or industries which have received financial contributions from their respective governments (European Commission, 2019). For example, one country exporting steel may have a steel industry which has received grants, tax credits and government purchases. These could be perceived by the home country as unfair competition against its own steel industry. Lastly, SPS measures are policies adopted by governments to stop the spread of diseases or pests among plants and animals. Regulations on live products that enter a country are essential to controlling and protecting their domestic markets from deadly infections. There are two main categories of SPS measures: Product Characteristic (PC) and Conformity Assessment (CA) measures, these categories define what aspects of the product the SPS is focussed on.

ADP	Anti-Dumping Policies
CV	Countervailing (or anti-subsidy) measures
SPS (PC)	Sanitary and Phytosanitary measures (Product Characteristic type)
SPS (CA)	Sanitary and Phytosanitary measures (Conformity Assessment type)

Table 1: *Main Types of Non-Tariff Measures (NTM) considered in this study.*

ADP and CV measures are applied bilaterally as an attempt to counter perceived unfair trade, in order to protect domestic industry. A study from the UN concluded that there are too many loopholes and ambiguities in the regulation of the use of ADP and CV measures, and this leaves room for them to be abused (Neufeld, 2001). The misuse of these measures can benefit the imposing country by protecting against competition, through additional duties on imports. Due to its effect on imports, if one country implements an ADP the opposing country would be incentivised to retaliate with an ADP back.

The literature suggests that the use of ADP (and to a lesser extent CV) may change the use of ADP or CV from opposing countries.

Literature has shown that SPS measures can significantly reduce trade between a country pair. They too have the potential to be misused as protection from fair foreign competition, but this has not been documented in the literature. This paper will investigate the gap in the literature through looking at the use of SPS and how they may interact with NTM use by the opposing country.

This paper aims to investigate the actual responses to imposed SPS measures by looking at the effect of each measure on the NTM (ADP and CV) use by the country the SPS measure was imposed on. This study will also investigate whether PC and CA types of SPS measures have different effects on an opposing country's NTM use. The difference in impact of groups of SPS measures within a short time frame versus individual SPS measures will also be investigated. This paper will not be looking at the impact of SPS on SPS measures as the method cannot be applied to this scenario (this is explained in section 3.2)

The rest of the paper is set out as follows; Section 2 provides an introduction to the different types of measures using established literature, Section 3 details the methodology and the specification of the model, Section 4 outlines the data sources, Section 5 contains the regression tables explains their results, Section 6 concludes the findings and finally, Section 7 gives the references.

2. Literature Review

The literature review will first look at the literature for each type of NTM (ADP, CV and SPS) one by one, explaining the difference between the two main subtypes of SPS (CA and PC), and then section 2.4 will focus on literature investigating interactions between NTM, looking at how NTM use by one can change NTM use by another.

2.1 Anti-Dumping Policy Literature

Anti-Dumping Policy (ADP) is viewed as one of the most important instruments in trade policy (Niels, 2000) (Davis, 2009). Most ADPs are used by the EU, USA, Canada and Australia, however in recent years there has been a significant rise in the use of ADP in emerging economies, especially China. Despite this, ADPs still cover a relatively small proportion of trade. During the period 1995 – 2010, ADP influenced between 1 and 3% of trade for each individual country (DFID, 2012).

The WTO established the rules for ADP in the General Agreement on Tariffs and Trade (WTO, 1986), outlining the specific conditions for a scenario to be recognised as dumping by the WTO. They require that the price of the exported good must be lower than the domestic price. Or, in the absence of a domestic price, comparison to the highest price available from a comparison country or to the cost of production of the product, plus a reasonable addition for a profit margin. The ADP then is a duty set on the import of products.

This WTO criteria is subject to much criticism. Niels (2000) summarises the empirical literature on the subject and states the main criticisms with the GATT conditions. Until regulation was changed by the WTO, conditions for price comparison were fairly relaxed and economies like the EU and US would use weighted averages in their favour, arbitrarily choosing which transactions to include in the measurement. Furthermore, sometimes the prices of different products are compared – low quality imports with high quality products domestically. A problem that is still seen today concerns the choice of comparison country. When a country is considered a planned economy, their prices cannot be used for comparison so another similar country is used. Mexico have used the USA as a substitute for China and Trinidad and Tobago as a substitute for Russia. The definition of planned economy can be loose.

Part of the GATT involves countries adding their own profit margins to cost levels, they would add an arbitrary cost of minimum 10% for overhead costs on top of production costs and an arbitrary 8% profit margin, in order to set the duty. This was ruled out by the WTO in 1994 and it stated that countries must justify cost increases by using details of the costs as evidence. However, this is still exploited by listing labour costs when in reality it is “family labour” i.e. fake costs or the inclusion of costs which actually had nothing to do with the product. (Niels, 2000).

Davis (2009) analyses the ADP use in the EU. The author discusses that ADP is at the heart of EU trade policy. Its use is based on the idea of establishing fair trade, although the author notes that economists are of the opinion that ADP has little to do with fair trade. The author finds that the main targets of such policy are exporters in growing markets particularly in emerging economies. ADP from the EU tends to be in declining industries which need help to stay competitive[1]. Furthermore, the author notes that ADP proposals must be submitted to the WTO for consideration before it can be certified, however once it has been initiated, the most likely outcome is that it will be passed through.

[1] This fits the findings of Stegmann (1991) who suggests that in many cases, the domestic producer is often less efficient that the importer who is being questioned, something that is not taken into consideration under antidumping law.

If there are loopholes so that ADP can be misused, then countries could use ADP effectively whenever they want to attack or respond to another country's use of trade barriers in order to protect their own industries.

ADP can have significant impacts on trade. Sandkamp (2018) finds that ADP use in the EU dampens trade and this lasts over time and also indirectly effects non-targeted exporters. Vandenbussche and Zanardi (2010) find that ADP causes large distortions, they find that users of ADP can cause an annual reductions of global imports by 5.9%..

2.2 CV literature

The WTO Agreement on Subsidies and Countervailing measures outlines the rules regarding subsidy use and measures to offset subsidies. It says that in order to implement a CV measure, a member must determine that there are subsidised imports, injury to a domestic industry and a causal link between the two.

CV measures are less common than ADP. Looking at data from the United Nations TRAINS database (United Nations, 2019) on almost all NTM across the world, the ratio of reports of ADP to CV is around 6 to 1. However, despite this they are still popular measures and they are still considered a major NTM (DFID,2012).

Chandra (2016) investigates the impact of US ADP and CV on Chinese exports and finds that both types of measures deflect Chinese exports to other countries by 15%.

Neufeld (2001) notes that ADP and CV have become popular substitutes for traditional trade barriers however, it is particularly anti-dumping policy that has been used for protectionist purposes. Misuse of CV is not commonly documented in empirical literature.

2.3.1 SPS literature

As explained in the introduction, SPS measures serve a valuable purpose, they help keep up standards in food safety and prevent the spread of disease. Just as the regulations of ADP enforcement are abused, the same could happen to SPS measures. There are many technicalities regarding inspections, labelling and licencing and it is possible countries can find ways to bend the rules to implement SPS measures in order to hamper trade. These measures are a big component of trade policy and as such there is a sizeable literature investigating the effects of SPS on trade volumes.

Recent empirical literature has highlighted the importance of SPS measures. The use of these measures has become increasingly more common, as can be seen in figure 1. This is a graph from the WTO which shows that the number of SPS measures implemented has increased from 198 implemented in 1995 to 1762 implemented in 2019. This shows that SPS measures are becoming increasingly more common and potentially more significant to research empirically.

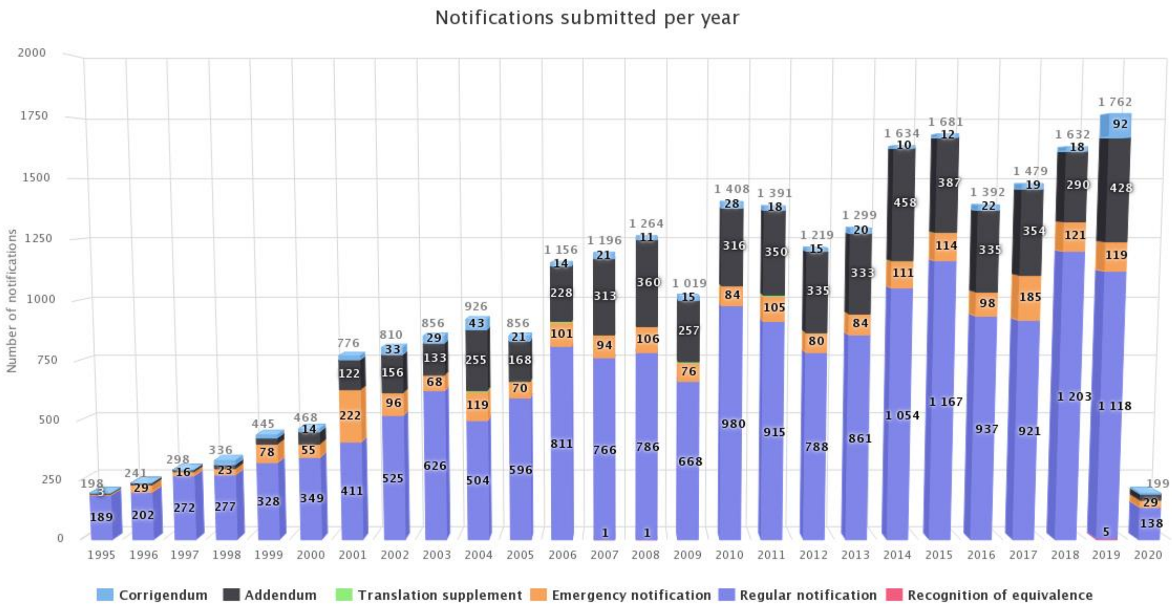

Figure 1: *This graph from the WTO IMS database shows the number of SPS measures submitted every year. The grey numbers on top of each bar represent the total number of SPS measures implemented in a given year, it is not a cumulative total. (WTO, 2019a).*

The empirical work covers a wide range of areas, such as the effect on specific products, on all products, between different regions and the effect of different types of SPS measures. The work tends to use the gravity model of bilateral trade to estimate the impact of different trade barriers on trade flows. This model was first introduced by Tinbergen (1962), developed further by Linnemann (1966) and also by Anderson and Wincoop (2003) along with many other authors. This idea adapts the model of gravity from Physics to explain that the volume of trade between two countries can be predicted by the distance between them, the size of their economies and a set of predictor variables. This type of estimation is one of the more robust models in international economics and it is widely used across the trade literature. It appears in many papers such as Khan et al. (2003) and Martínez-Zarzoso and Nowak-Lehman (2003).

Otsuki, Wilson and Sewadeh (2001) adopt a narrow focus and look at only one SPS measure; the maximum aflatoxin level in imports of 5 categories of food (at a HS2 aggregation[2]) by 15 European countries on imports from 9 sub-Saharan African countries. They estimate their model using some of the standard gravity model variables: distance, colonial ties and the GNP of the importing and the exporting country. They found that a 10% decrease in the maximum allowable level of aflatoxin in

[2] Harmonised System (HS) classification refers to the level of product aggregation it ranges from 2 to 6, 6 being the most disaggregated. This system gives each product line its own unique code.

Europe would lead to a 11% decrease in cereal exports and a 3.4% decrease in the exports of fruits, nuts and vegetables for sub-Saharan African countries.

Otsuki and Wilson (2002) look at the impact of pesticide residue limits on banana trade towards 12 countries from 19 developing countries and find a significant negative impact of the increase in this measure on the trade of bananas.

Jayasinghe et al. (2010) look at the impact of SPS regulations on the demand by foreign corn producers for US corn seed from 48 countries between 1989 and 2004 and find that they have a significant negative impact on US corn seed exports. The authors use a gravity model and Maximum Likelihood estimation with a Heckman correction. These methods are used to correct for the zero trade flows issues from bilateral trade data which cause problems in OLS estimation.

There are many more papers investigating SPS measures as a whole such as; Fontagné et al. (2005), Disdier et al. (2008), Kang and Ramizo (2017). These papers all find that SPS measures lead to negative effects on trade, Disdier et al. (2008) conclude that although SPS measures bring value in terms of consumer safety, they can be detrimental to trade for developing economies. Kang and Ramizo (2017) find that the impacts of SPS measures vary by region. They find that SPS measures hurt exports of developing Asian countries significantly whereas exports from non-Asian, non-OECD developing countries are boosted by SPS. They also find that intra-regional trade among developing Asian countries is reduced by SPS.

2.3.2 Main types of SPS: Conformity Assessment (CA) and Product Characteristics (PC)

Crivelli and Gröschl (2016) look at the impact of different types of SPS measures on trade in agricultural and food products across more than 150 importer and 150 exporter countries at a HS4 product level disaggregation between 1996 and 2010. They split SPS measures into two separate categories, conformity assessment measures (certificate requirements, testing, inspection and approval procedures) and product characteristics (pesticide requirements, quarantine requirements and labelling). The authors use data from the WTO Information Management System (IMS) which provides information on concerns from one country about an SPS measure implemented by another country. At first the authors estimate the model using SPS measures as a whole and they find that the probability of entering an export market is 16% lower if their SPS frequency measure increases by one unit. They then look at the separate effects of SPS measures and find that conformity assessment measures reduce the probability of trading bilaterally by 31% while concerns related to product characteristics have no significant impact on the likelihood of trade. Conformity assessment measures are also shown to have a *negative* and significant impact on the volume of trade whereas product characteristics have *a positive* and significant impact on the volume of trade.The explanation of why conformity assessment measures constitute a market entry barrier could be due to the relatively high costs and difficult procedures that are imposed on the exporter. This is because measures related to testing, inspection and approval are relatively more expensive than measures such as labelling. Product characteristic measures may have a positive result because the information content of these type of measures signals quality and safety which is a positive indication to consumers and possibly this benefit outweighs the cost of implementing the measure. The authors use the usual gravity model estimation with a maximum likelihood technique.

Fassarella et al. (2011) analyse the impact of SPS and TBT on Brazilian exports of poultry meat from 1996 to 2009 to the world's main meat importers. They find that the type of measure employed by a government can have very different impacts on trade. If the Brazilian government employed conformity assessment type SPS measures of inspections and licences this could lead to a reduction in trade, but if they implemented measures related to quarantine and labelling this could increase their trade volumes with the world. The authors use fixed effects and a Poisson pseudo-maximum likelihood estimation.

2.4 NTM Interaction Literature

This section looks at literature which investigates interactions between NTM. Interaction is defined as how the use of NTM by one country affects the use of NTM by an opposing country. Does the imposition of an NTM cause a reaction by the recipient country? Is there a responding increase in countermeasures? Or a conciliatory decrease? In general, there are not many published empirical studies of the interaction between a country applying NTM and responding change in NTM use from the second country. Some have looked at the effects of ADP measures or a composite of NTM measures on the response from the receiving country. No papers could be found investigating the effects of imposing SPS their interaction with other NTM. There is a gap in the literature to investigate SPS interactions and that was the inspiration for this study.

Blonigen et al. (2001) investigate the use of ADP in the US. The authors look at whether the use of ADP measures from the US on country X is affected by the threat of ADP retaliation from country X. The authors find that the ADP authority in the US is influenced by the threat of retaliation and it rejects petitions for ADP against countries which tend to feature in trade disputes. The authors suggest that the threat of retaliation from opposing countries is one of the reasons why ADP covers such a small proportion of trade. Feinberg and Olson (2004) look at ADP use from 41 importing countries who have filed at least one ADP measure and find that these countries' ADP use leads to retaliation by the opposing country. They use a probit binary choice model with fixed effects. The authors have access to a large sample of sector level ADP filings (over 400,000 observations) which they use to investigate the retaliatory component of ADP measures. With the industry level data they look at how firms in one sector may implement an ADP one year and investigate whether the opposing country has implemented one first in the previous year and whether it was to the same sector or not. They find that the likelihood of a country filing a case to a country who had imposed one on it the previous year is 7% greater than if the opposing country had not.

A paper by Gawande (1995) looks at whether US NTM are retaliatory, using a tobit model. The paper investigates whether US NTM are used partially to deter foreign protectionism. The idea comes from a game theoretic model by Baldwin (1990). Baldwin's model suggests that the threat of retaliatory measures can discourage a foreign country from implementing their own trade barrier in the first place. Gawande uses a HS4 product level disaggregation and focuses on nine countries, including some of the US' major trading partners (e.g. UK, Japan, the measure used for NTM is a composite measure including multiple types such as ADP and CV but not SPS. The author finds that US NTM contain a significant retaliatory component.US NTMs are shown to respond to NTM imposed by European countries on US exports of processed foods. One major drawback of this study is that retaliation to an NTM is only considered within the same product group. The US may respond in many ways, they are not limited to retaliating with the same or similar measure that was imposed on them, they may impose a measure on a different product group. Finally, this result does not tell us anything specifically about any individual NTM , or indeed about SPS in particular, which leaves room for further research.

In the literature described above there have been some investigations of the interactions of ADP and, to a lesser extent CV, but not with SPS. This study will investigate this gap, looking interactions of SPS measures with ADP, CV and with other SPS measures from opposing countries. To do this, it will check if the imposition of SPS measures, whether a single measure or a sequence of multiple measures, results in a change in the levels of NTM (ADP and CV). It will look at the impact of SPS measures as a whole and will also delve deeper to show the different impact of both individual types (PC and CA) of SPS. Previous literature has suggested that both PC and CA can constitute as a market entry barrier therefore they both could potentially be misused and the imposition of these measures could influence other countries NTM use.

Finally, the studies discussed here mainly focus on the US. This study will look at all countries which have ever raised on been subject to a WTO concern. This spans 79 countries, from this the countries features the most are Argentina, Australia, Brazil, Canada, China, European Union, India, Japan and the United States

3. Methodology

3.1 Specification

The main challenge in this study is how to demonstrate the effect of the introduction of an SPS measure on the use of the NTM by the responding country. Each database containing the NTM (details in section 4 below), shows start and end dates of each measure and the country pairs involved. Across the time frame, 1995 – 2018 there are many NTM used within a country pair, one country often repeatedly implements measures onto the same opposing country. To gauge the effect of the introduction of an SPS it has to be shown that NTM use is different before and after the SPS imposition. Using the start dates, one can compare the number of NTM implemented in a period before an SPS and the number of NTM occurring after it.

Looking at an example which takes two countries, country ***i*** and ***j***. In order to measure the effect of country ***i*** imposing an SPS measure on country ***j***, the change in NTM use by country ***j*** on country ***i*** needs to measured. This will show whether the SPS measure increased or decreased the number of NTM implemented. To capture these changes, there will be a “before” window and an “after” window, which is a period of time of length ***t*** which will count the number of NTM implemented from country ***j*** in each. Each SPS will have “before” and “after” windows created, on either side of its start date. Each SPS has its specific start date and therefore its own time windows, which will be the same length ***t*** for all SPS measures in one regression and then ***t*** will be varied between 1 and 15 months, in separate regressions, and in each case the results will be examined for statistical significance. This method will only count measures which have start dates that are *within* the time window, the length of the NTM are not considered here and measures which start before a window starts and overlap into the window, are not considered.

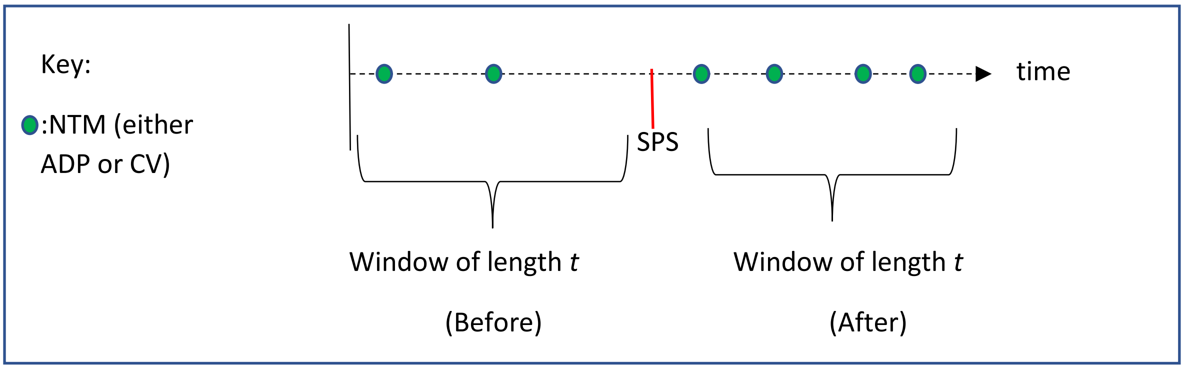

Figure 2: A visual representation of the SPS windows. It shows how the change in the rate of a specific measure is captured by set periods before and after an event.

As can be seen in Figure 2, the number of NTM (ADP and CV) either side of the SPS are captured in two time windows of length t. In this example, if *t* takes a value of six months, two NTM from *j* on *i* occurred in the six month period leading up to an SPS and four occurred in the six months after. Two different variables labelled *NTMafter* and *NTMbefore* record the values of NTM in each respective window, which would take values of four and six respectively.

Therefore, the rate of NTM implemented during a period of time before and after an SPS occurs will be compared.

The regressions in this study will be structured around the following form:

$$NTMafter_{jit} = \beta_0 + \beta_1 NTMbefore_{jit} + \varepsilon_{ijt} \qquad \textbf{\textit{(i)}}$$

$NTMafter_{jit}$: This records the number of NTM implemented by *j* on *i*, in period *t* after the date which an SPS from country *i* on *j* occurred.

$NTMbefore_{jit}$: This records the number of NTM implemented by *j* on *i*, in period *t* before the date which an SPS from country *i* on *j* occurred.

β_0: The intercept coefficient.

ε_{ijt}**:** This is the error term which contains unobserved factors.

The study will therefore regress the number of NTM (ADP and CV) from country *j* on *i*, occurring period *t* before an SPS from country *i* on *j*, against the number of NTM from *j* on *i* occurring period *t* after the

SPS from *i* on j. This is shown in formula (i). The aim of this is to look at a period *t* either side of an SPS occurring, and investigate the change caused by the SPS has on the number of NTM implemented per period *t* after, compared to before.

Each observation in the regression will represent a different SPS measure and then each SPS measure will have windows of length *t* before and after its start date. *NTMafter* and *NTMbefore* represent these windows. The value of these variables represents the number of NTM implemented during the length *t* of that window – for one type of NTM at a time. For example, in the six month window for the ADP regression, for SPS measure number #92 occurring on the 10/7/03, *NTMafter* will take a value two if there are two ADP initiated within six months of that date.

Coefficient β_1 shows the relationship between the number of NTM occurring before and after an SPS occurs, across all SPS. If β_1>1 and the result is statistically significant, this means that looking at all SPS, the number of ADP or CV implemented after them is greater than the number before. If β_1<1 and the result is statistically significant then that means that looking at all SPS, the number of NTM implemented after them is less than the number before.

Both scenarios have their own interpretations. If β_1>1 then this could mean that the SPS measure has sparked a retaliation from an opposing country in the form of increasing NTM. However, β_1<1 is also an interesting case, an interpretation of this could be that that the opposing country backed off as a result of the SPS and reduced the number of ADP or CV in force against the first country.

The method in this paper is preferable to a fixed effects panel data specification for several reasons. Panel data would not be helpful when trying to look at delayed interactions. It is fairly conceivable that countries take time to respond to NTM from opposing countries in a time frame of weeks or months. It would be hard to capture a relationship with panel data between things that occur at different times, especially when the time between them is different for each observation. Even if this was resolved, there are more problems. If the NTM were recorded as panel data, longer measures would mean more positive observations recorded over time. In a standard fixed effects type regression, looking at an SPS triggering another NTM, if the NTM lasts longer it would appear the SPS had a stronger effect on the implementation of other NTM. For this paper what is interesting is the number of NTM implemented as a result of an SPS, not the length of the NTM or SPS.

3.2 Investigating the impact of SPS measures on SPS measure use

Previous literature has shown the interactions between one country's use of ADP and opposing countries' ADP use. It makes sense to apply this scenario to SPS, and to investigate how the use of SPS measures can influence other countries to change their use of SPS measures. This was considered as part of this study and several attempts were tried to include it however, using the current method, which works well for analysing ADP and CV, one cannot accurately measure the effect on SPS use by SPS measures.

As explained in the last section, the method in this paper involves counting the number of NTM occurring before and after an SPS and measuring the change. It does not make sense to use this same method here i.e. to investigate the change in the "after" window relative to the "before" window when using SPS on SPS only. The reason this method works with ADP and CV and not SPS is because in ADP and CV the method involves comparing the SPS measures to another dataset whereas in this scenario there is only one set of information, the SPS list of measures.

This can be best demonstrated by examples. In the simplest case when analysing the impact of country **i**'s SPS on country **j**, a scenario like this could emerge:

Date of SPS measure	Imposing Country	Receiving country	NTMbefore	NTMafter
1/3/18	i	j	0	1
1/6/18	j	i	1	0

Table 4: *Example 1, showing SPS measures being implemented by country i on j and j on i in retaliation.*

In this scenario country ***i*** implements a measure first and then country ***j*** responds. When looking at the effect of country ***i*** on ***j*** it appears as though ***i*** has caused a reaction from country **j** because *NTMafter > NTMbefore* for country ***i***. However, when looking at the effect of country ***j*** on ***i***, it appears as though it has reduced SPS use from ***i*** because *NTMafter < NTMbefore.* The problem is that the interaction needs to be investigated from the point of view of both sides, country ***i***'s move and country ***j***'s move.

Taking a more complicated example helps explain this further:

Date of SPS measure	Imposing country	Receiving country	NTMbefore	NTMafter
1/3/18	*j*	i	0	1
1/4/18	*j*	i	0	1
1/6/18	*i*	j	2	1
1/9/18	*j*	i	1	0

Table 5: *Example 2, showing a sequence of SPS measures from country j on i and i on j.*

This Table shows a scenario where country ***j*** imposes two measures, country ***i*** imposes one and then country ***j*** imposes one back. From the point of view of country ***i***'s SPS, it appears as though country ***i*** has reduced SPS use by country ***j***, from two measures in the given period before to one measure after. However, from the first two SPS from country ***j***, it appears they have sparked an SPS from country ***i*** and then the final measure appears to have reduced SPS use. As can be seen here when looking at SPS against SPS it becomes confusing, these measures are dependent on each other and so this cannot easily be extended to a regression model. Measures will seem to be cancelling out one another and it will be hard to make any sense of the regression coefficients given. For this reason this method cannot be applied to investigate the impact of SPS on SPS measures.

In order to do this, one would need more data, perhaps the notifications database (described in more detail in section 4.1) which was not used in this study. One could compare the effect of SPS concerns on SPS notifications and it would not have this issue.

3.3 Conditions for Identification

Two aspects of the method are important to consider to gain a good identification in terms of the validity of the regression result:

1) **Window length :** The appropriate length of the windows is unknown and it is possible that different lengths of window have different outcomes. Some lengths of time may be too short for a response to take place, if it is too long then will contain other NTM which may not be related to the SPS. The length of time *t* will be varied monthly up to 15 months length in time.[3] This is to check that any one month window is not driving the results. This should not be the case anyway, as there is no reason a longer window would pick up more measures on one side than the other.

2) **Overlapping SPS measures:** If there are other SPS measures within the window of an SPS measure then it will be difficult to single out the "before" and "after" period of each without them affecting each other. For example if there is an SPS_{ij} in the "after" period of an SPS_{ij} , potential NTM implemented as a result of the first measure will be in the "before" window of the next SPS. This could bias the result, increasing the value of *NTMbefore*. The next section will outline the solution to this issue.

3.4 Overlapping SPS

When there are SPS measures that fall into a window of a different SPS measure (both from the same country to the same recipient), it can be hard to attribute individual NTM responses to any one SPS

[3] Not all months are reported in the results.

measure. In figure 3, there are two scenarios as examples, scenario A and scenario B and each SPS measure from ***i*** on ***j*** is labelled from one to six. Looking at SPS three in Scenario A, there are three NTM, from ***j*** on ***i*** , occurring before it and then only one after. It appears as though the SPS measure may be reducing the number of NTM. However, as can be seen there are two other SPS which could be considered to be influencing this change in NTM.

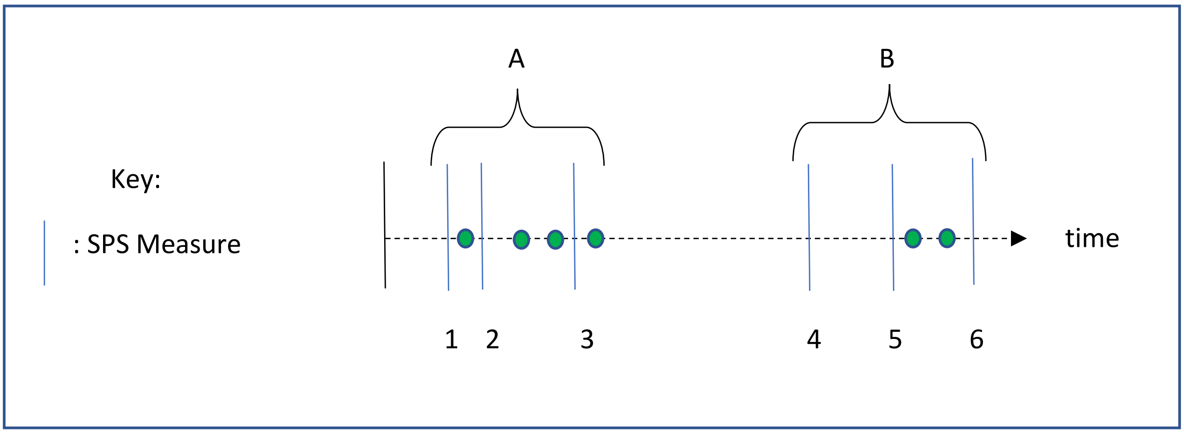

Figure 3: *This shows SPS measures grouping together. Each blue line is an SPS, each green dot is whichever measure is focussed on and each section is grouped into A group or B group.*

In order to address this issue, SPS_{ij} that occur within close approximation to other SPS_{ij} will be grouped together:

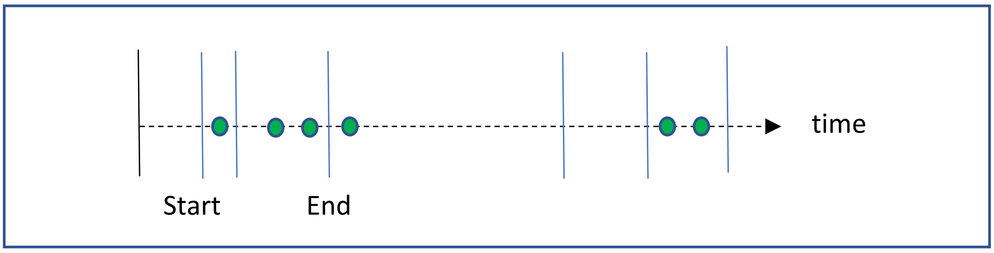

Figure 4: *The start and end of each SPS group will be identified. The NTM that occur in between these two measures will be added to the After variable for the group observation.*

The start and end of a group of SPS will be identified, the whole group will be condensed to one SPS measure: the end SPS. The NTM that occur within this group, will be added to the NTM that occur in the window after the last SPS occurs. For example in the scenario on the left, there are three measures occurring in between the start and the end, this will be added to the after window for the end SPS. The total *NTMafter* value will then be four. This removes any problems with overlapping SPS measures. The SPS measures with start dates close together can be thought of as one synchronised interaction from one country to another and therefore they can be condensed to one value. These groups will be created and then added to the database with the rest of the SPS measures. Different groups will be made up of different numbers of SPS measures, groups with more SPS measures are more likely to have a stronger effect than smaller groups. This will be investigated in the results section.

Only SPS measures *which have the same country **i** and **j** and occur within a short time of each other* will be grouped. It is hard to say exactly what the right time frame is between two SPS measures so that they can be grouped together. Different lengths of time will be tried to create the groups, and the results will be compared.

Data

4.1 SPS data

Data on SPS measures from 1995 to 2018 will be taken from the World Trade Organisation (WTO) Technical Barriers to Trade Information Management System (IMS) (WTO, 2019a). This database provides information on SPS specific trade concerns (STCs) and on notifications. When a country ***i*** imposes an SPS measure on country ***j*** country ***i*** must notify the WTO and this is recorded as a notification in the system. If country ***j*** has a problem with an SPS measure imposed by country ***i*** then it can raise this with the WTO and this is recorded as an STC[4]. This paper focuses on concerns rather than notifications, the STC database contains around 600 STCs. Using STCs rather than notifications gives the grounding for a stronger identification, it means that every positive observation will represent a scenario where one country has a problem with the SPS measure raised by the other and will be a potential misuse. These scenarios are much more likely to lead to retaliation than a standard SPS measure and therefore are of interest to this study. Furthermore, the notifications database over the same period contains over 20,000 SPS measures and each measure does not have the bilateral country related to it, shown in the documents available online. One would have to individually research each measure and find the bilateral country related, this is beyond the scope of this study.

The SPS IMS contains all measures which are raised in the SPS committee. The SPS committee are meetings between WTO members where members can raise concerns about SPS measures in accordance with the SPS Agreement. (WTO, 2019b). This agreement outlines the WTO rules on setting SPS standards in terms of requirements, harmonisation and transparency of measures.

The IMS database gives a list of all SPS concerns showing the reporting country(s) and the maintaining country(s)[5], the year of the concern and whether it has been resolved or not. The downloadable spreadsheet from the database website does not define STC's in terms of PC or CA, and so they need to be separated into their respective categories. In order to do this, each STC will be individually

[4] In this scenario country *i* is the raising country and *j* is the maintaining country.
5 The database considers the European Union as one country.

searched online in the WTO. In the online portal STCs can be searched, the database provides a short summary of the details of the concern and the nature of the measure. The descriptions will be compared with the definitions of conformity assessment and product characteristics measures taken from the WTO agreement Series[6] (WTO, 2019b). Furthermore, the spreadsheet only contains the years of each measure, precise dates for each measure will be taken from the online portal and added to the spreadsheet so that a precise comparison can be made.

4.2 NTM data

Data on ADP and CV measures were taken from the World Bank Temporary Trade Barriers database (WTTB, 2019). This contains data on NTM from the 1980s until 2015. Data from 2015 to 2018 was taken from the United Nations TRAINS database. This database gives bilateral data on many different types of NTM such as Anti-Dumping Policy (ADP), Quantity Controls (QC), Price controls (PC), Countervailing measures(CV) and Safeguards (SG). The available data for QC and PC measures were so small that they were not considered in this study. SG measures (measures are to protect domestic industry from potential increases in imports (WTO, 2019c)) are applied to all trade partners equally and are less interesting for this study.

As mentioned previously, ADP are measures to prevent or reduce dumping. Exporters from a country are said to be dumping when they export a product at a price lower than the price they normally charge in their home market. CV measures are also known as “anti-subsidy” measures. This is used in a scenario where another country has boosted their sector through subsidies and then exports into other markets. The WTO allows countries to apply counter taxes to stop these exports harming domestic sales.

The countries covered in this list are all countries which are part of the WTO, which includes most countries in the world. Despite this, the majority of ADP and CV measures are spread between emerging economies (India, China, Argentina, Brazil), the USA, Canada and the EU. These countries dominate the data with over 70% of all ADP and CV measures originating from one of these areas.

6 **Conformity Assessment**: procedures that are used to determine the relevant requirements in technical regulations or to make sure standards are fulfilled. They include procedures for certification requirements, sampling, testing, inspection, evaluation, verification of conformity, registration and approval.
Product Characteristics: concerns on requirements for process and production methods, transport, labelling, pesticide requirements, quarantine requirements and protected zones).

5. Results

The regressions consider the number of NTM occurring before and after an SPS, each observation refers to an SPS measure and each measure has a value for the number of NTM occurring before and after it occurs, within a specific time window. The time windows are varied between 1 and 15 months, as can be seen in the output tables.[7]

Regressions were run looking into the following areas:

1) the effect of SPS measures as a whole on different NTM individually (ADP and CV).

2) The effects of different types of SPS measure (PC and CA) on different NTM individually.

3) The effect of including group SPS measures (multiple SPS occurring in a short time frame of one another).

5.1 Proportion of SPS measures interacting with NTM

Before each different type of regression is observed, it is first important to look at how many SPS measures actually contain positive values for *NTMafter* or *NTMbefore*, or both. If an SPS measure contains a zero value for *NTMafter* and *NTMbefore* ,then this shows there was no interaction between SPS and NTM (ADP or CV).

Looking at Table 8 below, column 1 shows the proportions of SPS that contain positive values for either *NTMafter, NTMbefore* or both for ADP and CV. The proportions are approximately 30% and 5% respectively. These proportions are low and this suggests that overall , SPS measure use from one country does not affect the use of ADP and CV from another country. The vast majority of SPS measures (at least 70% for each), are not being used in response to previous NTM and did not trigger any NTM to occur after it.[8]

NTM	***NTMafter or NTmbefore or both >0***	**Ratio of PC measures to CA measures in total dataset**	**Ratio of PC to CA in the SPS with *NTMafter, NTMbefore or both >0***
ADP	28.7%	2.13	1.24
CV	5.12%	2.13	1.10

Table 8: *This table shows out of the total number of SPS measures and how many contained positive values for NTMafter, NTMbefore or both. The data for these tables comes from the ADP, CV and SPS datasets described in section 4.The ratios of PC to CA are averages across all time windows.*

[7] Regressions were run for months 1 to 15, however only 6,9,12,15 were reported, there were no significant advantages to including all the months, the trends can be seen by using the listed months.
[8] It is also possible that *NTMafter* and *NTMbefore* both have the same positive value this would suggest that there was no change as a result of the SPS measure. If this is the case then this will bring the coefficient value closer to one. A value of one would indicate the SPS had no impact on the NTM use.

Importantly, positive values for *NTMafter* or *NTMbefore* for one SPS measure is not enough to say that the SPS is interacting with the NTM, it is possible the "after" and "before" windows can contain NTM that are unrelated to the SPS measure. The regressions in the next sections will help show if these positive values are interactions or not through statistically significant coefficients for *NTMbefore*.

Why are these proportions so low (30 and 5%)? Previous literature has shown ADP can be very responsive to other countries use of ADP, why hasn't SPS shown to be the same? SPS measures are useful measures which ensure safety between countries so they are frequently used to do what they say they are for e.g. to ensure quarantines for potential diseases or ban imports of certain live animals. SPS measures are necessary for trade and it is inevitable to have some food safety regulations between countries. Considering then the wide use of SPS measures, it may then be expected that the number of measures that are used as a trade tool is a relatively low proportion. It is important to remember 28.7% of SPS have positive values for *NTMafter or NTMbefore,* when looking at ADP, which is a significant proportion, these measures will be explored further in the next few sections. CV measures, which interact to a lesser extent will also be explored.

What is interesting then is to give a closer look at the percentage of SPS concerns with positive values for *NTMafter or NTMbefore* or both, to check whether there is anything that might indicate why these concerns are different from the other 71.3%.

From the available data, there are three main areas that can be analysed; the dates of SPS measures, the countries involved and the types of SPS measures. Analysing the dates shown, the dates of these measures are no different to the remaining SPS and the countries featured are the typical selection of countries seen in the SPS database who use SPS the most : United States, China, European Union, India, Brazil, Argentina and Australia.

The interesting characteristic is the type of SPS involved. As shown in Table 8 (and by the difference in observation numbers between PC and CA regressions in section 5.2.2) the ratio of PC to CA SPS measures in the dataset is around 2.0, for every given time window. In general, throughout the data, the use of PC measures widely overshadows the use of CA measures. In the main SPS database, the number of PC measures is 410 and CA is 198. What is interesting then is that when looking at the percentage of SPS that have positive values for *NTMafter, NTMbefore or* both, the number of PC to CA reduces to a ratio close to 1, for both ADP and CV datasets. From this, one can conclude that for any given PC or CA SPS measure it is more likely that the CA measure will interact with another NTM than the PC measure. This fits the previous narrative explained in the literature review, that CA measures are generally considered more costly to trade and therefore are more likely to be used to hamper the trade of opposing countries. Furthermore, PC measures typically on quarantine and disease may be harder to misuse than CA measures on inspections and scheduled delays which may be less technical and easier to get away with misusing e.g. adding an extra inspection may be easier than issuing new quarantine rules.

From here on, the regressions shown will be looking at relationships between NTM for those proportions of SPS with a positive value for at least one of *NTMafter* or *NTMbefore*. A statistically significant result will help show that the SPS measures are related to the NTM values and show that these positive values are not simply random measures.

5.2 Interactions of SPS use with imposed levels of ADP and CV

5.2.1 Investigating the impact of SPS measures as a whole on ADP and CV use

This section focusses on Tables 9 and 11, where SPS measures are considered as a whole and SPS groups are not included.

VARIABLES	(1) 6 months	(2) 9 months	(3) 12 months	(4) 15 months
NTMbefore	0.680***	0.482***	0.803***	0.842***
	(0.042)	(0.024)	(0.032)	(0.032)
Constant	0.108***	0.094*	0.102	0.116
	(0.041)	(0.051)	(0.069)	(0.079)
Observations	608	520	505	489
R-squared	0.304	0.450	0.555	0.585

Table 9: *Regressing the number of ADP measures occurring before an SPS occurs, on the number of ADP occurring after, within windows of varying length. SPS are not separated into different categories. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1*

In reference to equation (i) in Section 3.1, β_1 is represented by the coefficient on *NTMbefore* in Table 9. A value of one would indicate that the *NTMbefore* and *NTMafter* are equal, a value higher than one would indicate that *NTMafter* is greater than *NTMbefore* and a value lower than one indicates *NTMbefore* is greater than *NTMafter.* In column 1, *NTMbefore* takes a value of 0.68 which suggests that the number of ADP measures implemented in the 6 months after SPS concerns are introduced are lower than in the 6 months leading up to it. This coefficient is significant to a 1% significance level. A value lower than 1 is seen for all different time windows shown from 6 to 15 months. This result covers all SPS concerns from 1995 to 2018. Referring again to Table 8, the proportion of SPS measures which had an ADP measure in its time windows is approximately 30%[9]. From this 30%, there seems to be a strong and significant result. The SPS measures cause a decrease in the number of ADP implemented per period after (for any given period after, from 6 – 15 months). The SPS implementation causes opposing countries to decrease their use of ADP, it appears as though the use of the SPS measure has caused them to back off. Why is this?

It is possible the SPS measure from country ***i*** has a drastic effect on the exports from the country it was imposed on, country ***j*** (this has been shown in previous literature, see section 2.3). As a result, country ***j*** highlights this SPS as a concern to the WTO and then reduces its use of ADP as it does not want to spark further measures from country ***i.***

The SPS measures in this case are all concerns, this gives a stronger identification because it increases the likelihood that the measure was used to hamper trade rather than help improve safety. This is signalled by the fact that the other country objected to it and therefore it is more likely that it was not strictly accurate or relevant. And this is more likely to be a misuse i.e. a situation where a country has used a measure to attack another country.

All time windows between 6 and 15 months showed similar results; the SPS measures reduced the implementation of ADP for each respective period.

Figure 5 shows scatter plots for ADP data behind regressions 2 and 3 for time windows 9 and 12 months respectively. From this graph it is hard to see the relationship between *NTMbefore* and *NTMafter* however, what can clearly be seen is that the vast majority of SPS measures contain values from 0 to 5 for *NTMafter* and *NTMbefore.* The blue dot on point (0,0) shows that many SPS have no positive values as discussed in section 5.1. There are SPS measures that contain value*s* reaching up to 10 and 15, indicated on the axes however the number of these is very low, as indicated by the colour scale.

[9] The number of observations shown in the Tables represents the total number of SPS in the dataset. This includes SPS measures that have zero values and are not in the regression. Therefore, the number of observations used in the regression is these observation totals multiplied by the percentage of non-zero SPS measures i.e. 28.7%.

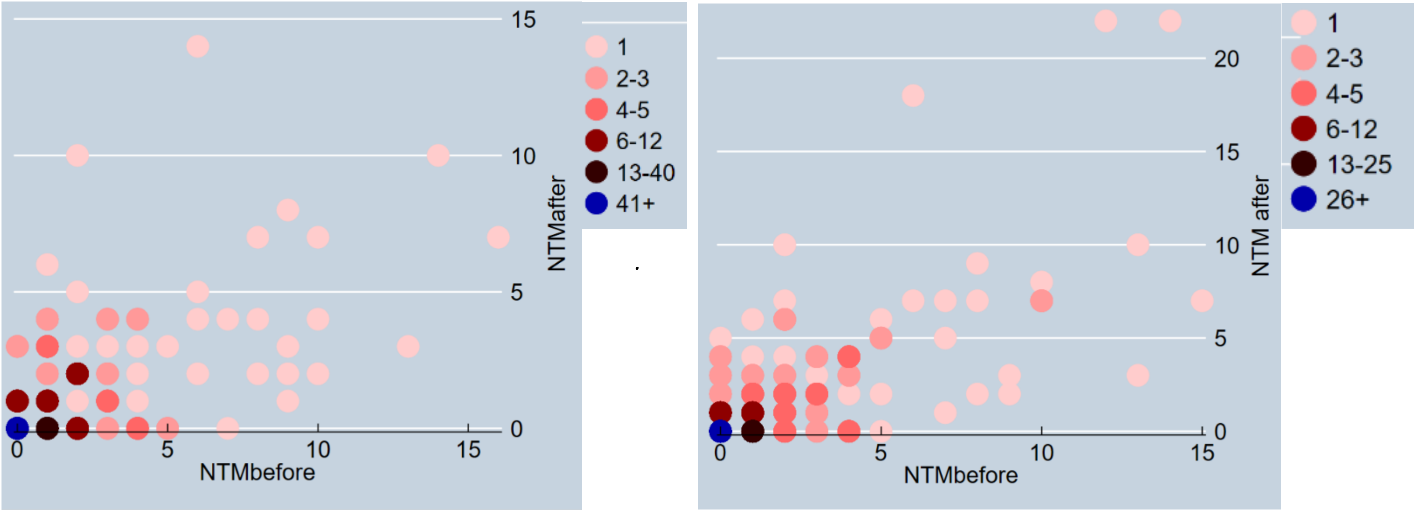

Figure 5: *Scatter plots showing the number of ADP 9 & 12 months (left and right respectively each legend is to the right of the respective graph) before and after each SPS occurs. Each point represents an SPS measure. The colour in the legend shows how many SPS values are at each point. Since there are many SPS which have the same numbers for NTMafter and NTMbefore, there are many SPS represented by one point on the graph. From Stata.*

In Table 11 below, when the interactions with SPS and CV are examined, the results show a different story to ADP. The coefficients across regressions 1 to 4 are all very close to a value of one. The value for 9 months is 0.941, 12 months is 1.033 and the other two are similarly close to one. This suggests that the implementation of an SPS measure has little effect on the use of CV measures by the opposing country. Some SPS may have seen an increase in CV and some may have seen a decrease however, overall they cancelled out leaving a coefficient around one suggesting that *NTMafter* is the same value as *NTMbefore.*

VARIABLES	(1) 6 months	(2) 9 months	(3) 12 months	(4) 15 months
NTMbefore	0.998***	0.941***	1.033***	0.998***
	(0.0342)	(0.0365)	(0.0328)	(0.0342)
Constant	0.00816	0.0103	0.00519	0.00816
	(0.0217)	(0.0154)	(0.0174)	(0.0217)
Observations	608	520	505	489
R-squared	0.629	0.522	0.663	0.629

Table 11: *Regressing the number of CV measures occurring before an SPS occurs, on the number of CV occurring after, within windows of varying length. SPS are not separated into different categories. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1*

The paper by Gawande (1995) suggested CV measure use from one country may change another country's use of the measure, but in general there is limited evidence to show this. Furthermore, misuse of CV is not well documented. Therefore, there is limited evidence to suggest CV measures interact with other country's use of NTM and the finding here is in line with this. In Table 8 of 5.12% and the regression coefficients being so close to one. There literature on CV is smaller than ADP, the ability of CV to influence other countries use of NTM and to impact trade is less documented and so the result that SPS does not affect CV as much as SPS affects ADP follows this pattern. However, all results from CV regressions must be met with caution because the number of SPS involved is so low. Only 5.12% of CV measures are involved, this means 5.12% of each number of observations given in each regression in Table 11.

5.2.2 Investigating the effect of different types of SPS measures (PC and CA) on NTM use

This section looks at differences in interaction from the different SPS categories; Product Characteristics and Conformity Assessment, on ADP and CV measures.

Starting with the ADP regressions in Table 12, the 6 month SPS window regression results in columns 1 and 2 show statistically significant values of 0.742 and 0.543 for PC and CA measures respectively. This means that CA measures have a stronger effect on reducing ADP measures than PC measures do, because the coefficient is smaller, meaning the SPS have a stronger effect on reducing ADP implementation. A similar result is found for the 9 month and 12 month windows. In the 15 month window, the opposite is found, the CA coefficient is larger than the PC one, however the other months show that there is a clear trend that PC coefficients are larger than CA coefficients.

VARIABLES	(1) 6 months [PC]	(2) 6 months [CA]	(3) 9 months [PC]	(4) 9 months [CA]	(5) 12 months [PC]	(6) 12 months [CA]	(7) 15 months [PC]	(8) 15 months [CA]
NTMbefore	0.742***	0.543***	0.482***	0.470***	0.640***	0.556***	0.709***	0.736***
	(0.040)	(0.091)	(0.020)	(0.052)	(0.024)	(0.069)	(0.030)	(0.075)
Constant	0.025	0.276**	0.019	0.137	0.060	0.250	0.067	0.204
	(0.035)	(0.107)	(0.040)	(0.125)	(0.043)	(0.154)	(0.063)	(0.169)
Observations	410	198	321	199	318	187	313	176
R-squared	0.460	0.159	0.653	0.361	0.691	0.306	0.644	0.406

Table 12: *Regressing the number of ADP measures occurring before an SPS occurs on the number of ADP occurring after, within windows of varying length. SPS are separated into different categories: Product Characteristics (PC) and Conformity Assessment (CA)[10]. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1*

Previous literature has suggested PC and CA measures both can limit market entry and this could explain why ADP measure use declines, the threat of even more SPS measures from country i may trigger country j to back off and reduce ADP use. In terms of the difference between PC and CA, CA measures can consist of inspections and licences which can be relatively more expensive to the exporter than PC measures. New inspections can cause delays to transport and license fees can increase trade costs. As a result CA measures may be considered more costly to trade, relative to PC measures, by the country it was placed on. Therefore, they may slow down with their ADP implementation as a result. In contrast, PC SPS requirements on quarantine and labelling are less likely to be used as a retaliation, however it is still possible.

[10] The respective definitions are reprinted here:
Conformity Assessment: procedures that are used to determine the relevant requirements in technical regulations or to make sure standards are fulfilled. They include procedures for certification requirements, sampling, testing, inspection, evaluation, verification of conformity, registration and approval.
Product Characteristics: concerns on requirements for process and production methods, transport, labelling, pesticide requirements, quarantine requirements and protected zones.

VARIABLES	(1) 6 months [PC]	(2) 6 months [CA]	(3) 9 months [PC]	(4) 9 months [CA]	(5) 12 months [PC]	(6) 12 months [CA]	(7) 15 months [PC]	(8) 15 months [CA]
NTMbefore	1.231***	0.631***	0.506***	0.565***	0.626***	0.622***	0.786***	0.717***
	(0.116)	(0.045)	(0.024)	(0.058)	(0.035)	(0.046)	(0.050)	(0.035)
Constant	0.028	0.029**	0.009	0.032	0.005	0.018	0.010	-0.002
	(0.021)	(0.013)	(0.009)	(0.023)	(0.013)	(0.025)	(0.024)	(0.025)
Observations	410	198	321	199	318	187	313	176
R-squared	0.239	0.520	0.543	0.364	0.508	0.552	0.445	0.752

*Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1*

Table 13: *Regressing the number of CV measures occurring before an SPS occurs on the number of CV occurring after, within windows of varying length. SPS are separated into different categories: Product Characteristics (PC) and Conformity Assessment (CA).*

The relevant regressions for CV measures can be found in Table 13. As mentioned previously the results for the CV regressions should be met with caution, due to the low proportion of SPS windows which capture CV measures. For the measures which do interact, in the 6,12 and 15 month windows, the PC measures have a higher coefficient than the CA measures. In the 6 month window the PC coefficient is 1.231 where as the CA coefficient is 0.631, this seems to suggest that implementing a PC measure increases opposing countries CV while CA measures decrease it. This is a strange result can be put down to low numbers of SPS measures in the regression. There is no real relationship to be seen and it cannot be concluded that one type of SPS is different to another when interacting with CV.

5.2.3 Investigating the impact of different numbers of SPS measures in a group on ADP and CV use

This section considers the SPS groups mentioned in section 3.4 and looks at the difference in NTM changes between groups with different numbers of SPS in them. *Group size* is a variable which takes a value according to the number of SPS represented in the observation. For individual SPS measures, which make up most of the observations, this takes a value of 1. When SPS are in close proximity and they are grouped together, *Group size* takes the value of the number of SPS measures making up that group. In the SPS dataset, the range of values for the size of the group is from 2 to 5.

Table 14 shows the same ADP regressions seen earlier in 5.2 only the variable *Group size* has been added. The *Group size* variable is statistically significant at the 1% level for the 9, 12 and 15 month windows. The 12 month window has a coefficient of 0.925 for the variable, this means that when the number of SPS measures in a group increases by 1, the number of ADP implemented after the group occurs, increases by 0.925. This is an interesting result, it suggests that when the number of SPS measures implemented from one country to another increases, the number of ADP introduced by the opposing country in return increases. This suggests a retaliatory component where the opposing country looks to counter the series of SPS "attacks" with an ADP in response.

VARIABLES	(1) 6 months	(2) 9 months	(3) 12 months	(4) 15 months
NTMbefore	0.677***	0.478***	0.758***	0.796***
	(0.042)	(0.024)	(0.032)	(0.032)
Group size	0.196	0.409***	0.925***	0.915***
	(0.184)	(0.138)	(0.146)	(0.157)
Constant	-0.094	-0.339**	-0.896***	-0.880***
	(0.194)	(0.155)	(0.171)	(0.187)
Observations	608	520	505	489
R-squared	0.305	0.460	0.588	0.611

Table 14: *Regressing the number of ADP measures occurring before an SPS occurs, on the number of ADP occurring after, within windows of varying length. The variable Group size has been included, this measures the effect of multiple SPS, in close proximity, on the ADP use of opposing countries. SPS are not separated into different categories. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1*

In 5.2 and in the regression in Table 14, implementing *one* SPS measure only leads to a decrease in the use of ADP afterwards, whereas increasing the number of SPS measures, within a close proximity to other SPS measures, increases the use of ADP afterwards. It is possible that when a country introduces one SPS measure, the other country backs off and does not want to spark further retaliation, but when a country introduces multiple SPS measures, the other country feels under threat and responds with ADP (and the more SPS placed on them, the more they respond).

Looking at the CV regressions in Table 15, The *Group size* variable is again statistically significant at the 9,12 and 15 month windows however looking at the size of the coefficients in each regression, the variable is economically insignificant. The coefficient in regression 3 suggests that if the number of SPS measures implemented on a country increases by one, then the CV from that country increases by 0.078. This is a very small number and it is relatively insignificant. This is unsurprising considering the analysis earlier, in section 5.1, that SPS measures seem to have little to no impact on the use of CV measures. Therefore, increasing the number of SPS measures beyond one i.e. increasing *Group size*, has an insignificant impact on CV use.

VARIABLES	(1) 6 months	(2) 9 months	(3) 12 months	(4) 15 months
NTMbefore	0.985***	0.937***	1.026***	0.993***
	(0.065)	(0.036)	(0.033)	(0.034)
Group size	0.004	0.060**	0.078**	0.077*
	(0.038)	(0.028)	(0.032)	(0.039)
Constant	0.023	-0.059	-0.083**	-0.079
	(0.044)	(0.036)	(0.041)	(0.050)
Observations	608	520	505	489
R-squared	0.282	0.526	0.667	0.632

Table 15: *Regressing the number of CV measures occurring before an SPS occurs, on the number of CV occurring after, within windows of varying length. The variable Group size has been included, this measures the effect of multiple SPS, in close proximity, on the CV use of opposing countries. SPS are not separated into different categories. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1*

There is one important point to note regarding the way the groups were constructed. As explained in section 3.4, SPS measures are grouped together when they are in close proximity to one another. There is a set limit on the length of time that is considered close proximity, which is relative to the length of the window. The length of time that allows for two SPS to be grouped together was set at t/2 – 1 e.g. for the 12 month window, this was 5 months. This was varied by a few months either side as a consistency check and there were no significant changes to the results.

However, given this method, some groups cover longer periods of time than others, as long as two measures are within the specific time frame from each other, they will be grouped together. This means that a group with 4 SPS measures could span a longer time than a group with 2 because the 4 could be slightly more spread out. It could be the other way around, that the 4 measures are all close together and the 2 measures spread over a longer period, however, it is probable that groups with more measures cover a longer period of time. This means that groups with more SPS measures are more likely to pickup more NTM in between the start and the end SPS measure, which is then added to the *NTMafter* value. Therefore, the *NTMafter* value is likely to be greater, the higher the number of SPS measures in a group, irrespective of the effect of SPS measure on other NTM use.

To investigate this further, the groups were analysed to look at group size (SPS quantity) and group length (in time). Across all month windows, over 90% of SPS groups are made up of measures that start on the same day. This includes groups even up to 4 SPS measures, all starting on the same day. This removes the threat that SPS groups with more measures cover longer periods of time which could undermine the result. The remaining SPS groups contain measures that do not start on the same day make up a small proportion of the groups (10%). Overall, this strengthens the result. It suggests that when countries implement several SPS measures on the same day rather than one measure, the opposing country responds by increasing its NTM use rather than reducing it. It suggests that in scenarios where one country implements many measures on another on the same day it triggers a different response to just one measure, and that the more that are implemented on one day, the stronger the retaliation in return.

6. Conclusion

Previous empirical literature shows that Non-Tariff Measures (NTM) such as Anti-Dumping Policy (ADP), Countervailing (CV) measures and Sanitary and Phytosanitary (SPS) measures can reduce trade volumes by reducing imports from the country they were imposed on. ADP has been widely misused across many countries to harm imports, and thereby to benefit the domestic economy. Furthermore, it has been shown that countries retaliate to each other's use of ADP with counter-ADP, and sometimes the threat of counter-ADP can limit other countries' readiness to use ADP. Some publications have looked at composite measures of NTM misuse, including ADP and CV usage, but so far SPS has not been studied for misuse or interactions. This paper asks if SPS may be misused in a similar way and can that influence another country's use of other NTM (in the form of ADP or CV)

The SPS concerns database was used as the primary data source in this study; it lists occurrences when SPS measures were highlighted as unfair by the countries they were imposed on. This increases the likelihood that any measure may be in place to harm the recipient country and a misuse has taken place.

This study uses a novel method which measures the number of NTM occurring before and after an SPS start date, within specific windows of time. It found that for the majority (almost 70%) of SPS measures, there was no interaction with ADP and CV measures. The explanation for this may be that since these measures are designed for a vital safety role, most of them will be put in place for their primary function and not to directly attack other countries.. However, around 30% of ADP measures and 5% of CV measures showed positive results relating to the implementation of SPS measures. Thus a sizeable proportion of SPS implemented generated a response in terms of ADP, changing the recipient country's use of NTM countermeasures and this indicates possible misuse. The fact that these SPS measures had been notified as concerns enhances their identification as possible misuse cases. The study found that these SPS measures reduced the number of received ADP measures implemented, relative to before. This was shown to be fairly consistent over a range of time periods studied from 1 – 15 months. This means that when these SPS measures were implemented, the opposing countries reduced their use of ADP measures in response. The SPS measures interacted with a smaller proportion (about 5%) of received CV and the results suggested that SPS had no effect on CV use in these proportions either, as well as the other 95%. In some cases, the CV changes that may have occurred for individual SPS measures cancelled out, showing no result. Investigating the 30% of SPS measures which potentially interact with ADP measures, a much larger proportion of them are Conformity Assessment (CA) measures than the proportion of CA measures in the entire database. This suggests that CA measures may be more likely to be misused to protect against foreign imports than PC measures. Furthermore, the regressions focussed on SPS type find that CA measures have a stronger effect on making opposing countries reduce their use of ADP and CV measures than Product Characteristic (PC) measures.

In the concerns database, SPS measures may be seen to occur either singly or in multiple events, sometimes spread over a period of time. When multiple SPS measures were implemented these were grouped together and analysed to see whether they created a reaction from the responding country. These groups contained several SPS measures that all start at similar times, many on the same day, aimed at the same opposing country. This study found implementing an extra 1 SPS measure within a short time of the first, increases the response of ADP in return by between 0.409 and 0.925. This suggested that when many SPS are applied it triggers a different response from the opposing country, compared to when only one SPS is applied. When one measure is applied, the opposing country backs off and reduces their ADP (but not CV) use, possibly to prevent further retaliation, but when several are applied at once, the opposing country increases their use of ADP (but not CV), possibly as they felt under threat from the multiple measures. The time windows of SPS measures and of the groups were varied to check whether this would impact the results however, they were shown to be robust to different length windows.

This study could not investigate the effect of SPS use on SPS measures due to nature of method and the data (explained in section 3.2). However, further work looking at using this method for not only the SPS concerns database from but also the notifications database from the WTO could prove worthwhile. There are potentially some interesting interactions found by looking at both SPS concerns and SPS notifications, from the WTO.

Further research could also investigate regional differences in effects of SPS measures, on the use of NTM by opposing countries. Research by Kang and Ramizo (2017) indicates that SPS measures have differing impacts on trade in different regions in the world. Furthermore, the proportion of NTM, including SPS, originating from emerging economies has increased significantly over the last decade. Future research could investigate these regional differences, in terms of income. It would be possible to categorise countries by their income level and compare the potential NTM use changes between countries of differing incomes e.g. an SPS from a high to a low income country, or from a medium to a high income country.

In order to ensure continuing low barriers to world trade it is important to learn more about these measures, how they may be abused, and how they interact. The WTO should tighten their rules on ADP, as it has been shown repeatedly throughout the literature that loopholes are regularly exploited. The WTO have implemented changes over the years to tighten regulation however, it is evident from literature that more could be done to prevent countries misusing this system. Misuse of ADP threatens competition and can harm productive industries, similarly misuse of SPS can harm exporters through large increases in export costs and by preventing market entry. This paper suggests that SPS measures, although primarily only used to ensure food safety, can be used to attack opposing countries, and that the Concerns Database is an important tool to monitor NTM abuse. Additionally, the WTO should review any regulations which allow SPS measures to be exploited, in order to protect fair trade and the welfare of exporters, particularly from developing countries.

7. References

Anderson, J. E., & Van Wincoop, E. (2003). Gravity with gravitas: a solution to the border puzzle. *American economic review*, *93*(1), 170-192.

Blonigen, B. A., & Bown, C. B. (2001) Antidumping and Retaliation Threats. *National Bureau of Economic Research.*(NBER Working Paper No. 8576). Retrieved from https://www.nber.org/papers/w8576.pdf

Chandra, P. (2016). Impact of temporary trade barriers: Evidence from China*. China Economic Review*, 38, 24-48.

Crivelli, P., & Gröschl, J. (2016). The impact of sanitary and phytosanitary measures on market entry and trade flows. *The World Economy*, *39*(3), 444-473.

DFID., (2019) *Anti-dumping measures and international trade: selected economic issues* [PDF file]. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/32460/12-754-anti-dumping.pdf

Disdier, A. C., Fontagné, L., & Mimouni, M. (2008). The impact of regulations on agricultural trade: evidence from the SPS and TBT agreements. *American Journal of Agricultural Economics*, *90*(2), 336-350.

European Commission (2019). The European Union and the United States reach an agreement on imports of hormone-free beef. Retrieved from https://europa.eu/rapid/press-release_IP-19-3012_en.htm

European Commission (2019).Actions against imports to the EU: Anti-Subsidy. Retrieved from https://ec.europa.eu/trade/policy/accessing-markets/trade-defence/actions-against-imports-into-the-eu/anti-subsidy/index_en.htm

Fassarella, L. M., Souza, M. J. P. D., & Burnquist, H. L. (2011). *Impact of sanitary and technical measures on Brazilian exports of poultry meat* (No. 321-2016-10878, pp. 1-19).

Feinberg, R. M., & Reynolds, K. M. (2006). The spread of antidumping regimes and the role of retaliation in filings. Southern Economic Journal, 877-890.

Fontagné, L., Mimouni, M., & Pasteels, J. M. (2005). *Estimating the impact of environmental SPS and TBT on international trade* (No. hal-00270511). HAL.

Gawande, K. (1995). Are US nontariff barriers retaliatory? An application of extreme bounds analysis in the Tobit model. The Review of Economics and Statistics, 677-688.

Jayasinghe, S., Beghin, J. C., & Moschini, G. (2010). Determinants of world demand for US corn seeds: the role of trade costs. *American Journal of Agricultural Economics*, *92*(4), 999-1010.

Kang, J. W., & Ramizo, D. M. (2017). Impact of sanitary and phytosanitary measures and technical barriers on international trade. *Journal of World Trade*, *51*(4), 539-573.

Khan, M, Chaudhary, Sunil & Xiumin, Li. (2018). The Factors Affecting Nepal's Trade: Gravity Model Analysis. European Academic Research. V. 6766-6782.

Linnemann, H. (1966). *An econometric study of international trade flows* (No. 42). North-Holland Pub. Co.

Martínez-Zarzoso, I., & Nowak-Lehmann, F. (2003). Augmented gravity model: An empirical application to Mercosur-European Union trade flows. *Journal of applied economics*, *6*(2), 291-316.

Maskus, K. E., & Wilson, J. S. (2001). *Quantifying the Impact of Technical Barriers to Trade: Can it be done?*. University of Michigan Press.

Neufeld, I. N., & CNUCED. (2001*). Anti-dumping and countervailing procedures: use or abuse?: implications for developing countries*. UN.

Niels, G. (2000). What is antidumping policy really about?. Journal of economic surveys, 14(4), 467-492.

Otsuki, T., Wilson, J. S., & Sewadeh, M. (2001). Saving two in a billion:: quantifying the trade effect of European food safety standards on African exports. *Food policy*, 26(5), 495-514.

Sandkamp, A. N. (2018). *The trade effects of antidumping duties: Evidence from the 2004 EU enlargement* (No. 261). Ifo Working Paper.

Stegemann, K. (1991). *The international regulation of dumping: Protection made too easy* (No. 2110-2018-3925).

United Nations. (2019). *The global database on Non – Tariff Measures* [Data Set]. Retrieved from https://trains.unctad.org/Forms/TableView.aspx?mode=modify&action=search

United Nations. (2020). What are Non-Tariff Measures?. Retrieved from: https://unctad.org/en/Pages/DITC/Trade-Analysis/Non-Tariff-Measures/What-are-NTMs.aspx

Vandenbussche, H., & Zanardi, M. (2010). The chilling trade effects of antidumping proliferation. *European Economic Review*, *54*(6), 760-777.

Wilson, J., & Otsuki, T. (2002). *To spray or not to spray? Pesticides, Banana Exports, and Food Safety*. The World Bank.

World Bank. (2019). *World Bank Temporary Trade Barriers Database* [Data set]. Retrieved from https://datacatalog.worldbank.org/dataset/temporary-trade-barriers-database-including-global-antidumping-database

WTO., (1986) *General Agreement on Tariffs and Trade* [PDF file]. Retrieved from https://www.wto.org/english/docs_e/legal_e/gatt47.pdf

WTO., (2019a). WTO Technical Barriers to Trade Information Management System. Retrieved from http://tbtims.wto.org/

WTO., (2019b). The WTO Agreements Series Technical Barriers to Trade. Retrieved from https://www.wto.org/english/res_e/publications_e/tbttotrade_e.pdf

WTO., (2019c). Safe guard measures . Retrieved from https://www.wto.org/english/tratop_e/safeg_e/safeg_e.htm