JAE Journal of Agricultural Economics

Journal of Agricultural Economics
doi: 10.1111/1477-9552.12316

The Effects of Non-tariff Measures on Agri-food Trade: A Review and Meta-analysis of Empirical Evidence

Fabio Gaetano Santeramo and Emilia Lamonaca[1]

(Original submitted May 2018, revision received September 2018, accepted November 2018.)

Abstract

The increasing policy interest and academic debates on non-tariff measures (NTMs) has stimulated a growing literature on how NTMs affect agri-food trade. The empirical literature provides contrasting and heterogeneous evidence, with some studies supporting the 'standards as catalysts' view, and others favouring the 'standards as barriers' explanation. To the extent that NTMs can influence trade, understanding the prevailing effect, and the motivations behind one effect or the other, is a pressing issue. We review a large body of empirical evidence on the effect of NTMs on agri-food trade and conduct a meta-analysis to disentangle potential determinants of heterogeneity in estimates. Our findings show the role played by the publication process and by study-specific assumptions. Some characteristics of the studies are correlated with positive significant estimates, others covary with negative significant estimates. Overall, we found that the effects of NTMs vary across types of NTM, proxies used for NTMs, and levels of detail of studies. The estimated effects are also influenced by methodological issues and publication processes.

Keywords: *Meta-analysis*; *non-tariff measures*; *trade barriers*; *trade standards*.

JEL classifications: *F13, F14, Q17, Q18*.

1. Introduction

While international trade negotiations have substantially reduced tariffs and fostered global trade, the level of non-tariff measures (NTMs) has increased over time. The growing use of NTMs has led to a less transparent trade policy environment (Fernandes *et al*., 2017), which calls for a deeper understanding of how NTMs influence trade.

---

[1]Fabio Gaetano Santeramo and Emilia Lamonaca are both with the University of Foggia, Italy. E-mail: fabio.santeramo@unifg.it for correspondence. The authors gratefully acknowledge support from the IATRC, and thank the anonymous referees and the JAE Editor, and John C. Beghin and the participants at the XV EAAE Congress and at the IATRC Annual Meeting 2017 for their valuable comments and suggestions on earlier drafts.

The importance of understanding the trade effects of NTMs in agri-food sector is attested by the large and increasing number of papers hosted in top journals (e.g. Hooker and Caswell, 1999; Wilson and Otsuki, 2004; Dawson, 2006; Sheldon, 2006; Olper and Raimondi, 2008; Cioffi *et al.*, 2011; Santeramo and Cioffi, 2012; Beckman and Arita, 2016). In particular, the number of studies on trade effects of NTMs has significantly increased since 2000 (14 studies available up to 2000, while 140 studies published up to 2017), an upward trend that is parallel to the upward trend in the number of implemented NTMs (from 1.09 million of NTMs notified in 2000, to 4.35 million in 2017[2]). However, the empirical evidence on their effects is heterogeneous. We examine this heterogeneity and, in particular, use meta-analysis to investigate the proximate 'causes', recognising that causality cannot be inferred from this approach.

These issues have been partly investigated by Cipollina and Salvatici (2008), Li and Beghin (2012), Beghin *et al.* (2015), and Salvatici *et al.* (2017). These reviews examine specific categories of NTMs or particular geographic areas: Cipollina and Salvatici (2008) pay attention to trade policies implemented at the border; Salvatici et al. (2017) focus on the trade measures of the European Union. In addition, the existing reviews are qualitative analyses, with the exception of Li and Beghin (2012) who also use a meta-analysis to investigate the variation in estimated effects of technical measures on trade of agri-food and manufacturing industries. By analysing a set of 27 papers on the gravity model, Li and Beghin (2012) show that some study characteristics (e.g. specific agri-food sectors, exclusion of multilateral resistance terms) are associated with trade-impeding effects of technical measures. Some issues, however, are still under-investigated. For instance, Li and Beghin (2012) do not include the effects of the review process (a major driver in meta-analyses) and the influence of type and proxy of NTM, as well as the aggregation level at which the study is conducted.[3]

Since Li and Beghin (2012), the number of studies of NTMs has grown substantially: in five years at least 49 papers have been published in peer-reviewed journals. This growth in the number of studies deserves further analysis. We use a newer, larger and updated dataset, and follow the methodological arguments of meta-analysis (Stanley and Jarrell, 1989; Stanley, 2005; Stanley *et al.*, 2008). In particular, we analyse a wide sample of 62 papers, which differ by type of measure investigated and theoretical framework (i.e. we do not restrict attention to gravity-based papers). In addition to the Sanitary and Phytosanitary Standards (SPS), Technical Barriers to Trade (TBT), and Maximum Residue Levels (MRL) which were the focus of Li and Beghin, we also include studies on additional standards with potential trade effects (e.g. protect consumers' health and safety, reduce asymmetric information) which include: quality and quantity control measures, private standards, voluntary standards, requirements on Genetically Modified Organisms (GMOs).

2. The Trade Effect of Non-tariff Measures

2.1. A theoretical perspective

Tariffs are protectionist by definition: they undermine the social welfare by crowding out trade (Swinnen, 2016). Non-tariff instruments may be protectionist or competitive

[2]UNCTAD (2017), The Global Database on Non-Tariff Measures, available at: trains.unctad.org (accessed 15 September 2017).

[3]Appendix S1 provides a detailed comparison with Li and Beghin (2012).

for trade: they imply welfare redistributions by addressing market imperfections such as asymmetric information and externalities (Xiong and Beghin, 2014). From a social perspective, while the optimal level of tariffs is zero, determining the optimal level of non-tariff instruments is challenging (Swinnen and Vandemoortele, 2011; Swinnen, 2017) due to the complex relationship linking trade and social effects of non-tariff instruments (Sheldon, 2012).

The understanding of non-tariff instruments has changed over time: as the term 'non-tariff barriers' (NTBs), which emphasises their protectionist scopes (e.g. quotas, export restraints), has been replaced by 'non-tariff measures' (NTMs), in order to capture their potential roles of either hampering or facilitating trade (Grant and Arita, 2017).

According to the definition proposed by UNCTAD (2012), NTMs are policy measures, other than ordinary customs tariffs, that may have economic effects on international trade of goods, changing traded quantities and/or prices. NTMs may also have a corrective role, by reducing asymmetric information (Technical Barriers to Trade, TBTs), mitigating risks in consumption, improving the sustainability of eco-systems (Sanitary and Phytosanitary Standards, SPSs), and influencing the competition and the decision to import or export (non-technical NTMs).

In a small open economy the policy-maker sets NTMs on a product category, produced in the domestic market and imported from the country's trading partners, in order to maximise domestic welfare: in a domestic market, the optimal level of NTMs depends on the trade-off between the marginal utility gain for consumers and the marginal cost for producers. The effects on domestic welfare are influenced by the trade strategies of trading partners. Unless the effects on domestic production exactly offset the effects on domestic consumption (Swinnen, 2016), NTMs influence trade.

From the consumers' perspective, NTMs are socially desirable and provide higher social well-being: by reducing asymmetric information and/or externalities, NTMs enhance consumers' trust, reduce transaction costs and increase consumers' demand (Xiong and Beghin, 2014). The growing demand and the higher costs of implementing NTMs increase the equilibrium price and, as a consequence, the consumption expenditures. The net effect of NTMs on consumers' surplus depends on the magnitude of (positive) utility gain compared to the size of (negative) effect on consumption expenditures: the higher the consumers' utility, the higher the willingness to pay a higher price for products under regulation (Crivelli and Gröschl, 2016; Swinnen, 2016).

From the producers' perspective, NTMs imply higher costs of compliance, both fixed costs (e.g. upgrade of practice codes and facilities, acquisition of certificates, conformity in marketing requirements) and variable costs (e.g. prolonged delivery time due to inspection and testing procedures at custom points, rejection of certain shipments, denial of entry of certain shipments) (Xiong and Beghin, 2014; Crivelli and Gröschl, 2016), resulting in a reduction in profits and supply. The reduced supply increases the equilibrium price and producers' revenue. The net effect on producers' profits depends on the magnitude of (positive) gain in revenue, compared to the size of (negative) implementation costs: the lower the implementation costs, the higher the gain in revenue for products under regulation (Swinnen, 2016).

For exporters, a NTM implemented in the destination country implies higher costs of compliance and a higher import price. If the difference between import price pre- and post-NTM is greater (smaller) than the difference between domestic price pre- and post-NTM, domestic producers face smaller (greater) implementation costs and obtain greater (lower) profits than foreign producers. The NTM acts as barrier (catalyst) for trade if it reduces (increases) domestic imports (Swinnen, 2017).

2.2. *An empirical perspective*

Several studies suggest that NTMs hamper trade (e.g. Peterson *et al*., 2013; Dal Bianco *et al*., 2016), others conclude that they foster trade (e.g. Cardamone, 2011), while numerous studies show mixed effects of NTMs on trade (e.g. Xiong and Beghin, 2011; Beckman and Arita, 2016). Such heterogeneity may reflect the different types of NTM, and their different rationales (Schlueter *et al*., 2009), but might also be a consequence of study design and methods.

Hardly any studies provide a general assessment of the trade effects of NTMs: a remarkable case is Hoekman and Nicita (2011) who suggest that NTMs are major Barriers to trade of agri-food products. Most empirical studies are more focused on specific NTM-, product-, or country-case studies.

Li and Beghin (2012) note that the lower the aggregation of NTMs under investigation, the crisper the policy implication for addressed issues. Technical Barriers to Trade (TBTs) tend to be catalysts for trade (e.g. de Frahan and Vancauteren, 2006), whereas Sanitary and Phytosanitary Standards (SPSs) show mixed evidence (e.g. Schlueter *et al*., 2009; Jayasinghe *et al*., 2010; Crivelli and Gröschl, 2016). Divergences may be due to the peculiarity of the SPSs, which may have either: 'a substantial positive impact [... or] a significant negative impact' (Schlueter *et al*., 2009, p. 1489). Maximum Residue Levels (MRLs) tend to act as barriers to trade (e.g. Otsuki *et al*., 2001a,b; Chen *et al*., 2008; Ferro *et al*., 2015).

The effects of NTMs may be also sector- and/or product- specific. For instance, NTMs appear to be trade-impeding for seafood products (e.g. Anders and Caswell, 2009), meat (Wilson and Otsuki, 2003), fruits and vegetables, cereals and oil seeds (e.g. Otsuki *et al*., 2001a,b). However, trade of fats and oils seems unaffected by beyond-the-border policies (e.g. Xiong and Beghin, 2011).

In addition, the countries involved in empirical analyses may generate specific geo-economic patterns of NTMs. Studies investigating the impacts of NTMs implemented by developed countries against developing countries are frequent in literature, and tend to show negative effects on the trade performances of developing countries (e.g. Anders and Caswell, 2009; Disdier and Marette, 2010). NTMs may have either negative (e.g. Yue and Beghin, 2009) or positive effects (de Frahan and Vancauteren, 2006) on trade between developed countries, though NTMs appear to limit trade among developing countries (Melo *et al*., 2014).

Other sources of heterogeneity may be related to the variety of methodological and empirical approaches we find in literature. For instance, different proxies are used to measure NTMs: inventory measures (e.g. dummy or count variables, frequency index, coverage ratio, prevalence score); computation of price gaps; estimated *ad valorem equivalents* (AVEs) (Gourdon, 2014). Our review suggests that the effects on trade tend to be negative if NTMs are proxied by AVE (e.g. Olper and Raimondi, 2008; Arita *et al*., 2017), or by frequency index and/or coverage ratio (e.g. Jongwanich, 2009; Fernandes *et al*., 2017). However, if

NTMs are proxied by dummy or count variables, the results may be either positive (e.g. Cardamone, 2011; Shepherd and Wilson, 2013) or negative (e.g. Peterson *et al.*, 2013; Dal Bianco *et al.*, 2016).

Lastly, different types of data matter. For instance, both Schlueter *et al.* (2009) and Beckman and Arita (2016) estimate the effect of SPSs on meat trade between developed countries. The former, using data aggregated at HS-4 digit, find a positive effect on trade while the latter use data aggregated at HS-6 digit and estimate a negative trade effect.

3. The Meta-analytical Approach

The empirical literature on trade effects of non-tariff measures (NTMs) is, perhaps, an archetypical example of the heterogeneity of the results of any economic research (Havránek, 2010). Since the pioneering work of Stanley and Jarrell (1989), economic meta-analyses (MAs) have been used to identify patterns in this heterogeneity.

Meta-analysis (MA) is an econometric approach that combines and summarises evidence from different but broadly comparable empirical studies, to illuminate the variation in results across studies (Stanley *et al.*, 2008), synthesising empirical estimates (Stanley and Doucouliagos, 2012). Doucouliagos and Stanley (2013) suggest that theoretical competition may shape the distribution of reported empirical findings and indicate publication selectivity. Publication selection has been a main concern of MA: it may bias estimates and disguise heterogeneity across studies, undermining the validity of inferences and policy implications (Stanley *et al.*, 2008; Doucouliagos and Stanley, 2013). Publication selection biases may concern the propensity of academics towards a particular direction of results (i.e. negative or positive estimates) (type I bias), or may occur if statistically significant results are treated more favourably, thus are more likely to be published (type II bias) (Stanley, 2005; Stanley and Doucouliagos, 2014).

However, MA is becoming more popular in economics, tending to focus on synthesis of results, for instance, for the price elasticity of demand (Böcker and Finger, 2017), the calorie-income elasticity (Santeramo and Shabnam, 2015), and to food safety effects (Xavier *et al.*, 2014). It has been also used to investigate international trade: Rose and Stanley (2005) and Havránek (2010) analyse the effect of currency unions on trade; Disdier and Head (2008) examine potential causes of variation in distance effect on bilateral trade; Cipollina and Salvatici (2010) investigate the impact of preferential trade agreements on intra-bloc trade; Li and Beghin (2012) explain variations in estimated trade effects of technical barriers to trade. Our focus is on trade effects of NTMs in the agri-food sector, following and updating the Li and Beghin (2012) study.

3.1. Literature searching criteria and selection process

There has been a substantial growth in NTM studies (theoretical, 32%, and empirical, 68%) during the last 25 years (see Figure A1 in Appendix S1).

We systematically reviewed the literature following the Stanley *et al.* (2013) guidelines. We searched studies in bibliographic databases (Scopus, Web of Science) while JSTOR provided access to multidisciplinary information from prestigious, high impact research journals. RePEc, IATRC, AgEcon Search and Google Scholar allowed us to cover grey literature[4] (i.e. working papers and conference proceedings); repositories of specific peer-reviewed journals[5] and papers' cross-references traced back further works. The search was carried out in August 2017 and was limited to research published in the period 1990 to 2017.

We used keywords such as 'trade' and 'agri-food trade', combined with other terms: 'non-tariff measure/non-tariff barrier', 'technical barrier to trade', 'sanitary and phytosanitary standard', 'maximum residue level', 'specific trade concern'. We identified 155 studies. Subsequently, each paper has been reviewed in depth, so to limit the analysis to papers that assess the trade effects of NTMs: we excluded theoretical studies, and papers that do not provide comparable empirical results.[6] The final sample includes 62 papers (47 published in peer-reviewed journals, and 15 from grey literature),[7] 1,362 observations (point estimates of trade effects of measures, ETEMs) and 1,213 estimated *t*-statistics.[8] Of these, 34 (56%) generated negative trade effects as their main result, while 21 (34%) found mixed results. Only 4 found a positive trade effect, while 3 found no significant effects.

3.2. Empirical model

Our empirical meta-analysis uses the *t*-statistics of ETEMs ($\hat{t}$) as the dependent variable, as recommended by Stanley (2001), to avoid heteroskedasticity, regressed on the precision of the estimates (i.e. the inverse of the estimated standard error, $\frac{1}{\hat{\sigma}}$), on J regressors related to the characteristics of the study (χ_j), and on K regressors related to potential publication selection (Z_k):

$$\hat{t} = \alpha_0 + \alpha_1 \frac{1}{\hat{\sigma}} + \sum\nolimits_{j=1}^{J} \beta_j \frac{\chi_j}{\hat{\sigma}} + \sum\nolimits_{k=1}^{K} \lambda_k Z_k + \varepsilon. \tag{1}$$

In line with Stanley *et al.* (2008), we include variables (χ_j) that are likely to influence the estimates, but that are uncorrelated with the likelihood of acceptance, as well as variables (Z_k) that may influence the likelihood of acceptance for publication, but should not be informative on the estimates. The constant term (α_0) collects potential information on the publication selection that is not directly included in the model

---

[4]In line with other analyses on trade issues (e.g. Disdier and Head, 2008; Cipollina and Salvatici, 2010; Li and Beghin, 2012), we include working papers and conference proceedings to identify publication selection. The journal prestige may be a source of publication bias (Santeramo and Shabnam, 2015). In order to avoid double counting, we include working papers and conference proceedings that do not correspond to revised versions published in peer-reviewed journals.

[5]We consider the following journals: *European Review of Agricultural Economics*, *American Journal of Agricultural Economics*, *Journal of Agricultural Economics*, *Agricultural Economics*, *Australian Journal of Agricultural and Resource Economics*, *Canadian Journal of Agricultural Economics*, *Applied Economics Policy Perspective*, *World Bank Economic Review*, *World Development*, *Agribusiness*, *Journal of Development Economics*, *Journal of Development Studies*, *China Agricultural Economic Review*, *German Journal of Agricultural Economics*.

[6]Appendix S2 provides a list of excluded studies.

[7]Appendix S3 provides Table A4 and descriptive statistics for each paper included in the sample (Table A5).

[8]We have 149 missing values for *t*-statistics due to the lack, in some papers, of standard errors and *t*-values.

(Stanley and Jarrell, 1989). The error term (ε_i) is assumed to be independently and identically distributed (i.i.d.).

We estimate equation (1) through a robust regression technique to mitigate potential problems related to outliers and influential data points (Belsley *et al*., 1980). Influential data points are likely to exist in our sample because we use multiple estimates from the same study (that are likely to be correlated). Coefficients of the robust regression allow us to infer the magnitude of ETEMs.

In order to determine which drivers may explain the direction (positive or negative) of statistically significant ETEMs, we use a Multinomial Logit (MNL) model: the dependent variable is categorical (Y_{MNL})[9] which classifies the ETEMs as negative (t-statistic lower than -1.96), not significant (t-statistic between -1.96 and 1.96), or positive (t-statistic higher than 1.96):

$$Y_{MNL} = \begin{cases} -1 \; \mathit{if} \; \hat{t} \leq -1.96 \\ 0 \; \mathit{if} \; -1.96 < \hat{t} < 1.96 \\ 1 \; \mathit{if} \; \hat{t} \geq 1.96 \end{cases} \tag{2}$$

By substituting the equation (2) into equation (1), we derive a system of two equations:

$$\begin{cases} \ln\left(\frac{Pr(Y_{MNL}=-1)}{Pr(Y_{MNL}=0)}\right) = \alpha_0 + \alpha_1 \frac{1}{\hat{\sigma}} + \sum_{j=1}^{J} \beta_j \frac{\chi_j}{\hat{\sigma}} + \sum_{k=1}^{K} \lambda_k Z_k + \varepsilon \\ \ln\left(\frac{Pr(Y_{MNL}=1)}{Pr(Y_{MNL}=0)}\right) = \alpha_0 + \alpha_1 \frac{1}{\hat{\sigma}} + \sum_{j=1}^{J} \beta_j \frac{\chi_j}{\hat{\sigma}} + \sum_{k=1}^{K} \lambda_k Z_k + \varepsilon \end{cases}, \tag{3}$$

where $\ln\left(\frac{Pr(Y_{MNL}=-1)}{Pr(Y_{MNL}=0)}\right)$ and $\ln\left(\frac{Pr(Y_{MNL}=1)}{Pr(Y_{MNL}=0)}\right)$ are the logarithms of the probability of generating, respectively, negative and significant or positive and significant ETEMs.

We use a Probit model to identify the drivers of statistical significance, and a Tobit model to investigate the magnitude of the estimated t-statistics (accuracy of ETEMs). The dependent variable of the Probit model is a dummy variable equal to 1 if the i-th ETEMs is statistically significant (t-statistics lower than -1.96, or higher than 1.96), and 0 otherwise:

$$Y_{Probit} = \begin{cases} 1 & \mathit{if} \; \hat{t}_i \leq -1.96 \; \mathit{or} \; \hat{t}_i \geq 1.96 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

The dependent variable of the Tobit model is a continuous variable equal to the t-statistics of ETEMs ($\hat{t}$), if it is larger than the threshold value (1.96 in absolute value), and 0 otherwise:

$$Y_{Tobit(-1.96)} = \begin{cases} \hat{t}_i & \mathit{if} \; \hat{t}_t \leq -1.96 \\ 0 & \mathit{if} \; \hat{t}_i > -1.96 \end{cases} \quad \text{and} \quad Y_{Tobit(1.96)} = \begin{cases} \hat{t}_i & \mathit{if} \; \hat{t}_i \geq 1.96 \\ 0 & \mathit{if} \; \hat{t}_i < 1.96 \end{cases}. \tag{5}$$

In the Tobit model, both right- and left-censored, positive coefficients imply greater t-statistics and, thus, less accurate ETEMs; *vice-versa* for negative coefficients.

In line with Li and Beghin (2012), we use a robust estimator of the clustered error structure to estimate MNL, Probit and Tobit models. We assume independence among clusters (i.e. papers), and dependence among observations within each cluster (i.e. ETEMs of the same paper).

---

[9]Differently from Li and Beghin (2012), we classify negative significant ETEMs as '–1' (instead of '1'), not significant ETEMs as '0' (instead of '2'), and positive significant ETEMs as '1' (instead of '3').

3.3. Description of covariates

Our model includes covariates related to the characteristics of the study (χ_j) and to the publication selection (Z_k), to account for heterogeneity in estimated effects of non-tariff measures (NTMs). The set of χ_j covariates controls for types of NTMs, proxies for NTMs, and the level of detail of the study. In particular, specific dummy variables account for types of measure (i.e. Technical Barrier to Trade, *TBT*, Sanitary and Phytosanitary Standard, *SPS*, Maximum Residue Level, *MRL*[10]).

Further dummies are used to proxy the intensive and the extensive margins of the NTM: *ad valorem equivalent*, or AVE, proxies the intensive margins by capturing the degree of protectionism of the NTMs (i.e. how much NTMs affect trade); dummy variables and indices proxy the extensive margins (e.g. existence or not of NTMs).

We also use dummy variables to identify the level of detail of the study. A dummy controls for the geo-economic affinity of countries that implement NTMs (i.e. reporters) and of countries affected by NTMs (i.e. partners): we classify reporters and partners into Northern (Developed Economies) and Southern (Developing Economies and Economies in transition) countries, according to the classification of the United Nations (United Nations, 2017). Other dummies control for the level of product aggregation (according to the Harmonised System[11]), and for the specific product category under investigation.

We control for potential publication selection: some covariates are related to methodological issues, others to the publication process. In particular, we control for the adoption of fixed effects to account for multilateral trade resistance terms in gravity models, and for the treatment of zero trade flows. The zero trade flows problem is a common issue in studies based on the gravity equation. Of our sample 87% are based on the gravity model, 3% rely on the Wales and Wooldland's Kuhn-Tucker model or the Monopolistic competition model, and 10% do not specify any model.

As for the publication process, we account for the prestige of the publication outlet, and for grey literature with specific dummies: one dummy controls for papers published in Q1 journals (according to the rank provided by the Scimago Journal & Country Rank at the date of publication for the subject area 'Economics and Econometrics') and one dummy accounts for working papers. Furthermore, we use a dummy variable to control for the presence of more than one article published by the same author. Table 1 lists our covariates.

[10]Countries frequently fix MRLs, as an alternative to SPSs, in order to ensure safe imports. The requirements on MRLs are not set in the WTO consultations, but they may be assimilated to the SPS A200 that sets the tolerance limits for residues and imposes a restricted use of certain substances in food and feed (UNCTAD, 2012). Due to these considerations, we distinguish SPSs and MRLs in separate categories.

[11]Commonly known as Harmonised System (HS), the Harmonised Commodity Description and Coding System is the internationally standardised system of names and codes used to classify traded products. We consider four levels of aggregation: HS-2 digit, which corresponds to a Chapter (e.g. 09 - Coffee, Tea, Mate and Spices), HS-4 digit, which corresponds to a Heading (e.g. 0901 - Coffee, whether or not roasted or decaffeinated; Coffee husks and skins; Coffee substitutes containing coffee), HS-6 digit, which corresponds to a Sub Heading (e.g. 090121 - Coffee, roasted, not decaffeinated), and HS-8 digit which corresponds to a Subheading determining duty (e.g. 09012100 - no distinction with respect to 090121).

3.3.1. Collinearity diagnostics

Our empirical model involves several dichotomous variables: potential collinearity may arise and confound estimation results. In order to test for collinearity, we use the variance inflation index (VIF) and the condition number, which should not exceed the threshold values 10 and 15, respectively, to avoid problems of multicollinearity (Belsley *et al.*, 1980).

As preliminary analysis, the correlation matrix shows that the correlation between the coefficients for 'Publication bias' and 'SPS', 'N-S' and 'SPS' for 'F&O' and 'Oilseed', and for 'Beverage' and 'Meat' is remarkably high (more than 0.80). Diagnostic outputs suggest possible strong collinearity between the same variables. We dropped the covariates with the relative higher VIF: 'Publication bias', 'Oilseed', 'N-S' and 'Beverage'. Collinearity diagnostics without the problematic covariates show no additional problems.

4. Results and Discussion

4.1. Description of the sample

Several studies contain multiple estimated trade effects of measures (ETEMs): in order to keep important information, we include all available evidence (Jeppesen *et al.*, 2002), rather than opting for the preferred estimate method, or for a synthesis of the estimates (Card and Krueger, 1995; Stanley, 2001; Rose and Stanley, 2005).

Table 2 summarises the ETEMs: 56% (759 point estimates) are negative; 44% (605 point estimates) are non-negative (positive or zero). A majority of estimates (61%) are statistically significant, with 508 point estimates (37%) being negative and 325 point estimates (24%) being positive. The mean and median values of (total) ETEMs are, respectively, –0.30 and –0.05 (confidence interval ranges from –3.33 to 2.73).

Figure 1 presents the distribution (boxplots) and the kernel densities of total, positive, and negative ETEMs[12]: half of the statistically significant ETEMs (680 observations out of 1,364) range between median values of (statistically significant) negative ($Me_{Neg.}$ = –0.42) and (statistically significant) positive ($Me_{Pos.}$ = 0.34) observations (Figure 2, panel (i)). The distribution of total ETEMs is bimodal with one negative peak ($Mo_{Neg.}$ = –0.21) and one positive peak ($Mo_{Pos.}$ = 0.02). Negative and positive ETEMs are almost equally dispersed (the standard deviations of negative, $\sigma_{Neg.}$ = 2.65, and positive ETEMs, $\sigma_{Pos.}$ = 2.97 are close) (Figure 2, panel (ii)).

Several variables seem correlated with the magnitude and the direction of NTMs' trade effects (e.g. geo-economic areas, type of NTMs). Figure 2 shows that the ETEMs differ across geo-economic areas. Papers that investigate North-North or North-South trade (40 papers, 822 point estimates) are larger in number than papers that analyse South-North or South-South trade (3 papers, 76 point estimates) (Figure 2, panel (i)). The variability of ETEMs across papers that consider North-North ($\sigma_{N\text{-}N}$ = 3.19) or North-South ($\sigma_{N\text{-}S}$ = 4.17) trade is larger compared to that of papers for South-South trade ($\sigma_{S\text{-}S}$ = 0.68). However, the median of ETEMs is lower for North-South ($Me_{N\text{-}S}$ = –0.10) trade than for North-North ($Me_{N\text{-}N}$ = –0.01) trade (Figure 3, panel (ii)).

[12]The distribution and kernel density estimated in Figure 2 refer to a subsample which ranges between the 10th and the 95th percentiles.

Table 1

Description of covariates and basic statistics

Covariates	Description	Type of variable	Detail	Mean
t-statistic	*t*-statistics related to ETEMs, estimated in literature	Continuous		–0.10
Standard error	Standard error related to ETEMs, estimated in literature	Continuous		0.91
TBT	Standard is a Technical Barrier to Trade	Dummy	1 if TBT (0 otherwise)	0.15
SPS	Standard is a Sanitary and Phytosanitary Standard	Dummy	1 if SPS (0 otherwise)	0.27
MRL	Standard is a Maximum Residue Level	Dummy	1 if MRL (0 otherwise)	0.18
AVE	*Ad valorem equivalent* (AVE) used to proxy standard	Dummy	1 if AVE (0 otherwise)	0.06
Dummy variable	Dummy variable used to proxy standard	Dummy	1 if dummy variable (0 otherwise)	0.35
Index	Frequency index or coverage ratio used to proxy standard	Dummy	1 if index (0 otherwise)	0.10
N-N	North-North	Dummy	1 if reporter and partner are developed countries (0 otherwise)	0.28
N-S	North-South	Dummy	1 if reporter is a developed country and partner is a developing country (0 otherwise)	0.32
HS-2 digit	Product aggregated at 2 digits of Harmonised System	Dummy	1 if product is aggregated at HS-2 digit (0 otherwise)	0.26
HS-4 digit	Product aggregated at 4 digits of Harmonised System	Dummy	1 if product is aggregated at HS-4 digit (0 otherwise)	0.36
Meat	Product category under investigation is meat	Dummy	1 if product category is meat (0 otherwise)	0.10
Dairy	Product category under investigation is dairy produce	Dummy	1 if product category is dairy produce (0 otherwise)	0.05
Cereal	Product category under investigation is cereals	Dummy	1 if product category is cereal (0 otherwise)	0.09
Oilseed	Product category under investigation is oil seeds and oleaginous fruits	Dummy	1 if product category is oilseeds (0 otherwise)	0.13
F&O	Product category under investigation is animal or vegetable fats and oils	Dummy	1 if product category is fats and oils (0 otherwise)	0.05

Table 1
(*Continued*)

Covariates	Description	Type of variable	Detail	Mean
Beverage	Product category under investigation is beverage	Dummy	1 if product category is beverage (0 otherwise)	0.04
Country-pair f.e.	Country-pair fixed effects	Dummy	1 if country-pair fixed effects are used (0 otherwise)	0.07
Time f.e.	Time fixed effects	Dummy	1 if time fixed effects are used (0 otherwise)	0.16
Product f.e.	Product/industry/sector fixed effects	Dummy	1 if product fixed effects are used (0 otherwise)	0.04
Zero trade	Treatment of zero trade flows	Dummy	1 if zero trade flows are treated (0 otherwise)	0.48
Q1	Peer-reviewed journal in the 1st quartile of Scimago Journal & Country Rank (SJR) at date of publication and subject area 'Economics and Econometrics'	Dummy	1 if published in Q1 (0 otherwise)	0.42
WP	Grey literature: working paper	Dummy	1 if published as WP (0 otherwise)	0.14
Authors	Paper co-authored by academics experienced on the issue	Dummy	1 if experienced author (0 otherwise)	0.61

Table 2
Descriptive statistics of the estimated trade effects of measures (ETEMs)

ETEMs	Min	Max	Median	Mean	SD	C.I.*	Obs.[†]
Total	−38.540	54.140	−0.048	−0.305	3.039	[−3.334; 2.734]	100%
Non-negative (positive or zero)	0.000	54.140	0.280	1.031	2.966	[−1.935; 3.997]	44%
Negative	−0.001	−38.540	−0.456	−1.37	2.653	[−4.023; 1.283]	56%
Significant	−38.540	18.105	−0.160	−0.598	3.151	[−3.749; 2.553]	61%
Significant non-negative (positive or zero)	0.000	18.105	0.580	1.263	2.296	[−1.033; 3.559]	24%
Significant negative	−38.540	−0.004	−0.734	−1.789	3.047	[−4.836; 1.258]	37%
Not significant	−12.920	54.140	0.004	0.155	2.795	[−2.640; 2.950]	39%

Notes: In the sample, only two observations are equal to zero.
*Confidence interval (C.I.) ranges between mean minus standard deviation (minimum) and mean plus standard deviation (maximum).
[†]Percentages computed on the total number of observations (1,364).

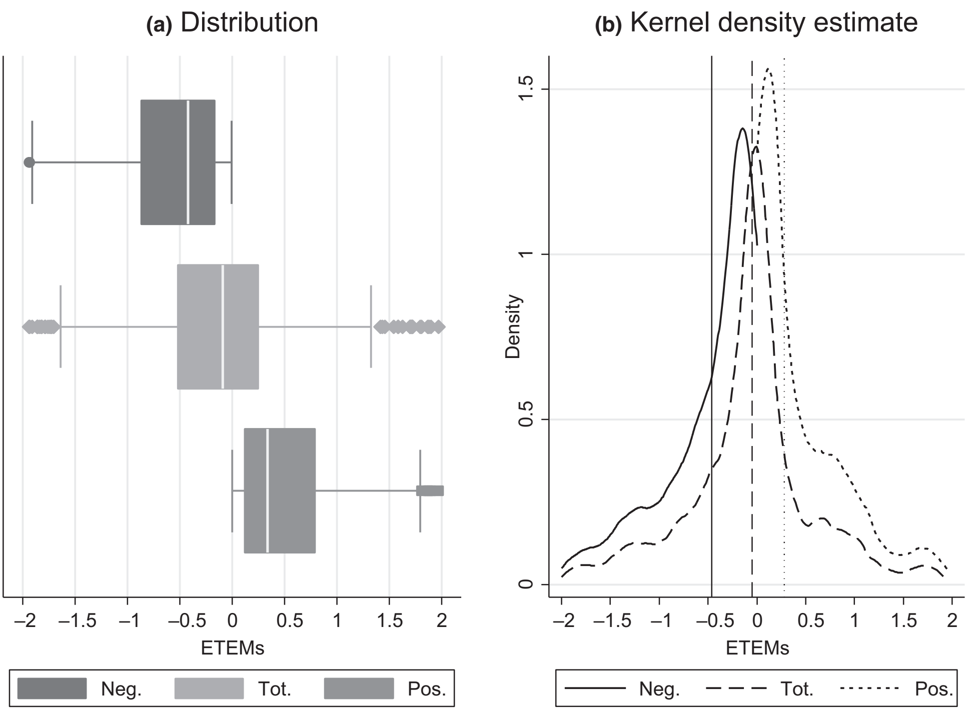

Figure 1. Estimated trade effect of measures (ETEMs) arranged by direction
Notes: In panel (a), distributions of ETEMs are on statistically significant observations within the 10th and the 95th percentiles. Horizontal lines within boxes are median values (Me) (i.e. $Me_{Neg.} = -0.42$, $Me_{Tot.} = -0.16$, $Me_{Pos.} = 0.34$). In panel (b), the estimated densities for ETEMs are computed removing observations which exceed the 10th and the 95th percentiles. Horizontal lines are median values (Me) computed on total observations (i.e. $Me_{Tot.} = -0.05$, $Me_{Pos.} = 0.28$, $Me_{Neg.} = -0.46$).

The ETEMs differ also by types of measure (Figure 3). A majority of papers focus on measures aiming at protecting human health (17 papers on Sanitary and Phytosanitary Standards, SPSs, and 25 papers on Maximum Residue Levels, MRLs: 504 point estimates), while several studies (15 papers, for 362 point estimates) report evidence for NTMs not involved in specific categories (such as Technical Barriers to Trade, TBTs, SPSs, MRLs, or Specific Trade Concerns, STCs): these cases (grouped under the tag 'Other') show a large heterogeneity in estimates, ranging from more than 50 to less than –15 (Figure 3, panel (i)). The variabilities of ETEMs for papers that analyse different types of NTMs are similar ($\sigma_{TBT} = 3.51$, $\sigma_{SPS} = 3.66$, $\sigma_{Other} = 3.65$), with an exception made for papers on MRLs, for which the variability is rather low ($\sigma_{MRL} = 2.20$). While the median values of ETEMs associated with TBTs and MRLs are close to zero ($Me_{TBT} = 0.01$, $Me_{MRL} = 0.02$), the median for SPSs is negative ($Me_{SPS} = –0.11$) (Figure 4, panel (ii)).

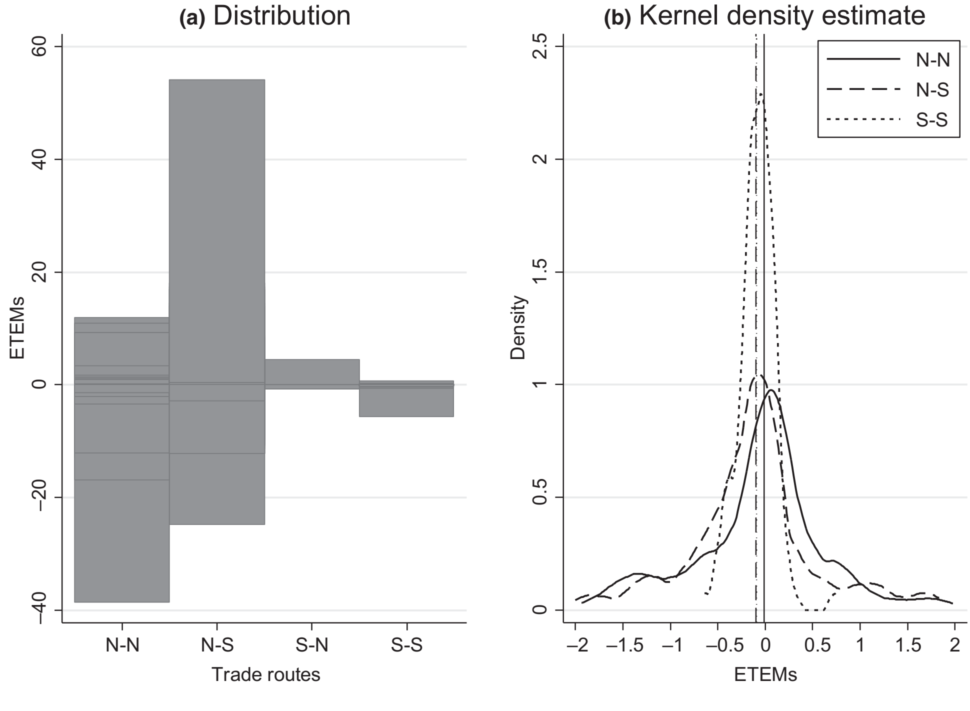

Figure 2. Estimated trade effect of measures (ETEMs) arranged by geo-economic area
Notes: Geo-economic areas are as follows: N-N stands for 'North-North', N-S stands for 'North-South', S-N stands for 'South-North', S-S stands for 'South-South', where the former are countries imposing NTMs (reporters) and the latter are countries affected by NTMs (partners). Reporters and partners are classified into North (Developed Economies) and South (Developing Economies and Economies in transition), according to the country classification proposed by the United Nations (2017). In panel (b), the estimated densities for ETEMs are computed removing observations which exceed the 10th and the 95th percentiles. Horizontal lines are median values (Me) computed on total observations (i.e. $Me_{Tot.} = -0.05$, $Me_{N\text{-}N} = -0.01$, $Me_{N\text{-}S} = -0.10$, $Me_{S\text{-}S} = -0.09$). Kernel density estimate for S-N is omitted because there are only two observations for ETEMs.

4.2. Regression results

We compare the results of robust regression, Multinomial Logit (MNL), Probit, and Tobit models.

The type of non-tariff measures (NTMs) clearly matters in determining magnitude, direction and accuracy of the estimated trade effects of measures (ETEMs). If a study focuses on Technical Barriers to Trade (TBTs) or Sanitary and Phytosanitary Standards (SPSs), ETEMs tend to be greater, more accurate if significant, and significantly negative with a lower probability (Table 3, columns A, B, E and F). ETEMs are greater also if Maximum Residue Level (MRL) is the measure under investigation, and are more likely to be significantly positive, but less accurate (Table 3, columns A, C and F).

As for the proxies used for NTMs, ETEMs tend to be lower and significant (Table 3, columns A and D). In particular, dummy variable and frequency index/

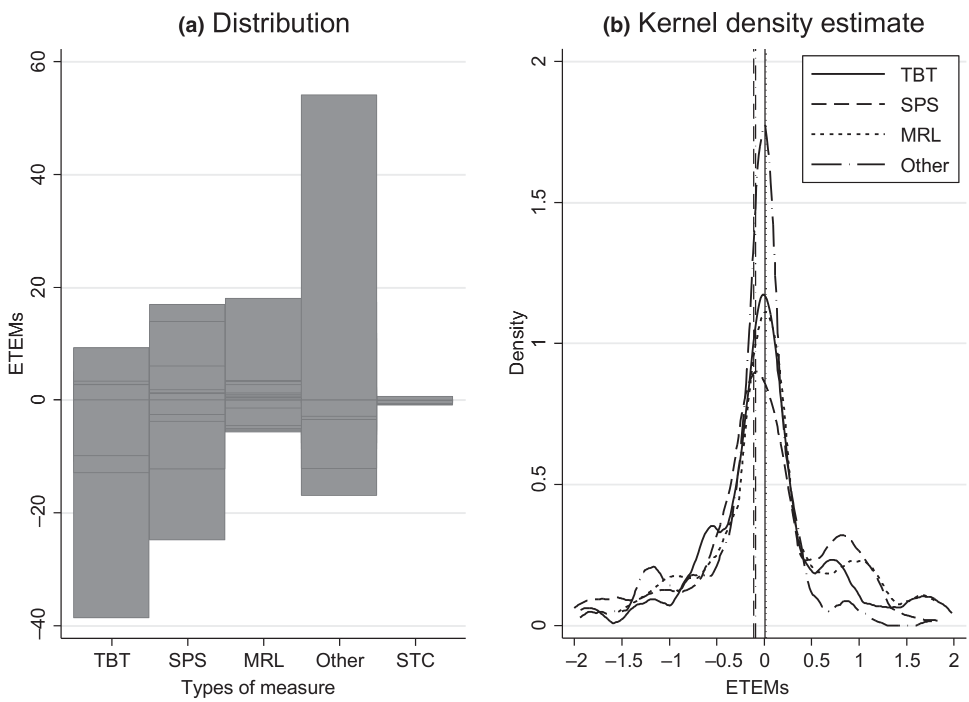

Figure 3. Estimated trade effect of measures (ETEMs) arranged by types of measure
Notes: Types of measure are as follows: TBT stands for 'Technical Barrier to Trade', SPS stands for 'Sanitary and Phytosanitary Standard', MRL stands for 'Maximum Residue Level', STC stands for 'Specific Trade Concern', 'Other' includes measures not involved in other categories (e.g. quality and quantity control measures, Hazard Analysis and Critical Control Points (HACCP), private standards, voluntary standards). In panel (b), the estimated densities for ETEMs are computed removing observations which exceed the 10th and the 95th percentiles. Horizontal lines are median values (Me) computed on total observations (i.e. $Me_{Tot.} = -0.05$, $Me_{TBT} = 0.01$, $Me_{SPS} = -0.11$, $Me_{MRL} = 0.02$, $Me_{Other} = -0.09$). Kernel density estimate for STC is omitted because there are only 24 observations for ETEMs.

Table 3
Regression results

Covariates	Robust regression $\hat{t}$ A	MNL $\ln\left(\frac{Pr(Y_{MNL}=-1)}{Pr(Y_{MNL}=0)}\right)$ B	MNL $\ln\left(\frac{Pr(Y_{MNL}=1)}{Pr(Y_{MNL}=0)}\right)$ C	Probit $Pr(Y_{probit} = 1 \mid B_{j,i}, \Gamma_{k,i})$ D	Tobit $Y_{Tobit(-1.96)}$ E	Tobit $Y_{Tobit(1.96)}$ F
β_1: TBT	0.118***	−0.180**	0.020	−0.016	12.170*	0.207
	(0.030)	(0.070)	(0.031)	(0.018)	(7.032)	(0.250)
β_2: SPS	0355*	−0.036***	3.04	−4.450	0.0002**	0.938**
	(0.211)	(0.334)	(4.82)	(3.070)	(0.008)	(0.389)
β_3: MRL	0.004***	−0.024	0.002**	0.001	2.031	0.020**
	(0.002)	(0.021)	(0.001)	(0.0004)	(1.689)	(0.008)
β_4: AVE	−0.161**	0.130***	0.091*	0.078***	−1.837	0.221
	(0.067)	(0.045)	(0.048)	(0.024)	(2.511)	(0.335)
β_5: Dummy variable	−0.130***	0.0619***	0.049	0.039***	−0.800	0.333
	(0.018)	(0.018)	(0.036)	(0.012)	(1.633)	(0.270)
β_6: Index	−0.056***	0.033*	−0.002	0.016*	−0.870	−0.069
	(0.011)	(0.020)	(0.026)	(0.008)	(0.566)	(0.153)
β_7: N-N	0.004	−0.001	0.009	0.004*	2.085***	0.116***
	(0.015)	(0.010)	(0.010)	(0.002)	(0.218)	(0.033)
β_8: HS-2 digit	0.098***	−0.026	−0.042	−0.016	−0.280	−0.241
	(0.030)	(0.042)	(0.031)	(0.014)	(2.604)	(0.303)
β_9: HS-4 digit	0.060***	0.007	0.007	0.002*	−1.904***	0.013
	(0.015)	(0.008)	(0.009)	(0.001)	(0.082)	(0.017)
β_{10}: Meat	0.119	0.397*	0.469*	0.275**	−15.680	1.342
	(0.139)	(0.223)	(0.265)	(0.117)	(11.730)	(1.687)
β_{11}: Dairy	−0.044	−0.183	−0.326	−0.147**	7.278	−3.143
	(0.088)	(0.112)	(0.238)	(0.063)	(8.311)	(1.937)
β_{12}: Cereal	−0.151*	−0.024	−0.295	−0.059	−0.221	−1.825
	(0.079)	(0.079)	(0.448)	(0.042)	(6.690)	(2.039)

Table 3
(*Continued*)

Covariates	Robust regression	MNL		Probit	Tobit	
	$\hat{t}$	$\ln\left(\frac{Pr(Y_{MNL}=-1)}{Pr(Y_{MNL}=0)}\right)$	$\ln\left(\frac{Pr(Y_{MNL}=1)}{Pr(Y_{MNL}=0)}\right)$	$Pr(Y_{probit} = 1 \mid B_{j,i}, \Gamma_{k,i})$	$Y_{Tobit(-1.96)}$	$Y_{Tobit(1.96)}$
	A	B	C	D	E	F
β_{13}: F&O	−0.048	0.059	0.063	0.031	−1.680	2.991***
	(0.042)	(0.060)	(0.045)	(0.027)	(1.671)	(0.013)
γ_1: Country-pair f.e.	−0.762**	1.561***	1.618**	0.948***	−39.160	8.520
	(0.311)	(0.558)	(0.813)	(0.284)	(43.530)	(7.368)
γ_2: Time f.e.	0.786	0.269	−0.113	0.091	−3.658	−4.339
	(0.567)	(0.523)	(0.478)	(0.277)	(27.740)	(4.674)
γ_3: Product f.e.	−1.285***	−0.008	−1.099*	−0.196	−16.010	−9.968*
	(0.239)	(0.325)	(0.631)	(0.195)	(22.260)	(6.010)
γ_4: Zero trade	−1.501***	0.672*	−0.019	0.258	−38.140	−4.422
	(0.278)	(0.404)	(0.432)	(0.182)	(27.060)	(4.836)
γ_5: Q1	−1.196***	0.070	−0.930**	−0.209	−28.950	−11.640**
	(0.357)	(0.443)	(0.408)	(0.179)	(25.440)	(5.842)
γ_6: WP	−0.718***	−0.722	−1.469**	−0.583	13.930	−8.198
	(0.267)	(0.576)	(0.739)	(0.358)	(28.760)	(7.864)
γ_7: Authors	−0.856*	1.335***	1.338*	0.812***	−50.060	7.062
	(0.455)	(0.447)	(0.719)	(0.221)	(39.960)	(6.401)
Constant	1.583***	−1.816***	−1.370*	−0.628***	176.400*	−14.790**
	(0.298)	(0.497)	(0.758)	(0.240)	(95.610)	(7.417)
Sigma					140.600**	19.020***
					(70.580)	(3.346)
Observations	1,210	1,213	1,213	1,213	1,213	1,213

Notes: Clustered standard errors are in parentheses.
***, ** and * indicate statistical significance at 1%, 5% and 10%, respectively.
The coefficient α_1 has been omitted because of collinearity.
The magnitude of estimated coefficients and related standard errors for variables 'SPS' are of the order of 10–15.

coverage ratio are more likely to provide significant positive ETEMs (Table 3, column B).

The level of detail of the study has a varying contribution on magnitude, significance, precision and direction of ETEMs. Studies that focus on NTMs across countries with similar levels of economic development ('N-N') tend to report ETEMs significantly different from zero (both positive and negative) and more accurate (Table 3, columns D, E and F). Our results expand on the findings provided by Li and Beghin (2012, p. 507), who observe that: 'when the NTM is a SPS policy regulating agri-food exports from a developing exporter to a developed importer, the probability to observe a trade impeding effect increases substantially'. Our results suggest that trade increasing effects are also found.

ETEMs are greater for data aggregated at the HS-2 digit or HS-4 digit levels (Table 3, column A): in particular, the higher the aggregation, the higher the probability of significant ETEMs, that tend to be less accurate if significantly negative (Table 3, columns D and E).

As for specific product categories, ETEMs tend to be greater for cereals (Table 3, column A), significant (either negative or positive) with higher probability for meat, but not for dairy (Table 3, columns B, C and D), more accurate if significant positive for fats and oils (Table 3, columns F). In line with Li and Beghin (2012) we show that technical measures are not likely to be trade-enhancing for processed food products (e.g. dairy produce, fats and oils), while we also found that this is not always true (e.g. meat).

Acronyms are as follows: North-North (N-N), North-South (N-S), Technical Barrier to Trade (TBT), Sanitary and Phytosanitary Standard (SPS), Maximum Residue Level (MRL), *ad valorem equivalent* (AVE), Harmonised System (HS), peer-reviewed journal ranked in the first quartile following the classification of Scimago Journal & Country Rank (SJR) at the date of publication and the subject area 'Economics and Econometrics'(Q1), working paper (WP).

Studies that include country-pair fixed effects generate lower and significant ETEMs (either negative or positive) with a higher probability (Table 3, columns A, B, C and D). Similarly, Li and Beghin (2012) pointed out that the trade effects of technical measures are influenced by the use of multilateral trade resistance terms. ETEMs are also lower in studies with time fixed effects (Table 3, column A), but significantly positive with a lower probability and, in these cases, less accurate in studies with product fixed effects (Table 3, columns C and F). On top of previous knowledge, we show that controlling for time and for product-specific (or sector/industry-specific) fixed effects impacts on magnitude, direction and accuracy of ETEMs.

If a study accounts for the treatment of zero trade flows, ETEMs tend to be lower and the likelihood of negative ETEMs significantly different from zero increases (Table 3, columns A and B). Similarly, Li and Beghin (2012, p. 507) argue that '*t*-values are more spread out in the negative range when zero trade is treated'.

Reflecting on the publication process, ETEMs are lower and less likely to be positively significant if provided in studies published in top journals (Q1) or in working paper series (Table 3, columns A and C). In particular, if significantly positive, ETEMs are less accurate if published in Q1 (Table 3, column F). In addition, ETEMs are lower and significantly different from zero (either positive or negative) with higher probability if a study is co-authored by experienced scholars (Table 3, columns A, B, C and D). Similarly, Havránek (2010, p. 254) argues that the authorship helps explain the direction and the magnitude of estimates.

Table 4

Summary of findings on the estimated trade effects of measures (ETEMs)

Covariates		Magnitude	Significance	Negative significance	Positive significance	Accuracy of negative significant ETEMs	Accuracy of positive significant ETEMs
Type of NTMs	TBT	Greater	n.s.	Less likely	n.s.	More accurate	n.s.
	SPS	Greater	n.s.	Less likely	n.s.	More accurate	More accurate
	MRL	Greater	n.s.	n.s.	More likely	More accurate	More accurate
Proxy for NTMs	AVE	Lower	More likely	More likely	More likely	n.s.	n.s.
	Dummy variable	Lower	More likely	More likely	n.s.	n.s.	n.s.
	Index	Lower	More likely	More likely	n.s.	n.s.	n.s.
Level of detail of the study	N-N	n.s.	More likely	n.s.	n.s.	n.s.	More accurate
	HS-2digit	Greater	n.s.	n.s.	n.s.	n.s.	n.s.
	HS-4digit	Greater	More likely	n.s.	n.s.	Less accurate	n.s.
	Meat	n.s.	More likely	More likely	More likely	n.s.	n.s.
	Dairy	n.s.	Less likely	n.s.	n.s.	n.s.	n.s.
	Cereal	Lower	n.s.	n.s.	n.s.	n.s.	n.s.
	F&O	n.s.	n.s.	n.s.	n.s.	n.s.	More accurate
Methodological issues	Country-pair f.e.	Lower	More likely	More likely	More likely	n.s.	n.s.
	Time f.e.	Lower	n.s.	n.s.	n.s.	n.s.	n.s.
	Product f.e.	n.s.	n.s.	n.s.	Less lilely	n.s.	Less accurate
	Zero trade	Lower	n.s.	More likely	n.s.	n.s.	n.s.
Publication process	Q1	Lower	n.s.	n.s.	Less likely	n.s.	Less accurate
	WP	Lower	n.s.	n.s.	Less likely	n.s.	n.s.
	Authors	Lower	More likely	More likely	More likely	n.s.	n.s.

Notes: n.s. stands for ‘not significant’; more/less are identified compared to the average effect.

Our analysis illuminates several issues: a number of variables contribute to explaining the heterogeneity in ETEMs. The magnitude of estimates is favoured by certain factors (type of NTM, product aggregation), but limited by other determinants (proxy for NTMs, 'cereal', 'country-pair f.e.', 'time f.e.', 'zero trade', publication vehicle). Some factors reduce the likelihood of having significant estimates ('dairy'), others increase this likelihood (proxy for NTMs, 'authors', 'country-pair f.e.', 'N-N', 'HS-4 digit', 'meat'). Moreover, some variables boost the probability of estimating trade-impeding effects (proxy for NTMs, 'authors', 'country-pair f.e.', 'zero trade', 'meat') and others hamper this probability ('SPS', 'TBT'). Similarly, the likelihood of estimating trade-enhancing effects may be either intensified ('authors', 'country-pair f.e.', 'MRL', 'AVE', 'meat') or limited ('Q1', 'WP', 'product f.e.') by specific variables. In addition, the accuracy of significant negative estimates increases with type of NTM and decreases with 'HS-4 digit'; *vice-versa*, the accuracy of significant positive estimates is favoured by certain variables ('SPS', 'MRL', 'N-N', 'F&O'), but not by others ('product f.e.', 'Q1'). Table 4 synthesises the evidence of our empirical models.

5. Conclusions and Policy Implications

The rapid growth of non-tariff measures (NTMs) has stimulated an interesting academic debate. Discriminating between the economics and the politics of NTMs is a challenge for academics and policymakers: theory suggests that NTMs may both stimulate and hinder trade (Swinnen, 2017). Accordingly, in literature, two opposite views prevail: 'standards as barrier' vs. 'standards as catalyst', with the empirical evidence being quite heterogeneous.

In order to characterise the heterogeneity in estimates, we qualitatively and quantitatively reviewed the empirical literature on the effects of NTMs on global agri-food trade. We explain the differences in findings, in terms of magnitude, direction, statistical significance and accuracy of estimates, with several control factors: types of NTM, proxies used for NTMs, level of detail of studies, methodological issues and publication vehicle. We build on the existing evidence provided in Li and Beghin (2012) with a considerably expanded dataset of estimates, and further details on methods.

We found that the estimated trade effects of measures (ETEMs) are greater for types of NTM, but lower for proxies of NTMs compared to the average effect. Maximum Residue Levels (MRLs) and *ad valorem equivalent* (AVE) tend to favour trade, whereas it is not always true that Sanitary and Phytosanitary Standards (SPSs) and Technical Barriers to Trade (TBTs) limit trade.

The level of detail also matters: for instance, analysing trade between developed countries or working with disaggregated data plays in favour of significant estimates. Magnitude, significance and accuracy of ETEMs may be also product-specific.

Last but not least, robust methodological approaches, and evidence provided by experienced authors, are correlated with greater chances of observing statistically significant, although lower ETEMs. Controls for specific methodological issues (e.g. inclusion of multilateral resistance terms, treatment of zero trade flows) seems to be advisable.

Our results highlight that magnitude, direction, statistical significance and accuracy of ETEMs are case-specific. This is in line with Livingston *et al.* (2008) who suggest that, in evaluating NTMs, economists try to compare benefits for trade and costs of management, production, market, and resource potentially related to an outbreak of disease or pest, finding differences on a case-by-case basis. Thus, the trade effects of

NTMs are likely to depend on specific countries, products and standards: generalisations are neither feasible nor sensible. A plausible explanation is that the variability in trade effects may reflect divergences among countries' food safety regulations and standards, differences in consumers' preferences across countries, ability (or limited capacity) to produce safe food, and willingness to pay for risk-reducing technology (Buzby and Mitchell, 2006; Jongwanich, 2009).

Our analysis illuminates the factors that appear to affect estimates of the effects of policy measures on trade, and illustrates how NTMs shape trade flows. The results also suggest that the proxy for NTMs matters, as does control for specific methodological issues, in order to provide more reliable estimates.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. A comparison with Li and Beghin (2012)

Figure A1. Trend of published researches on trade effects of NTMs over time.

Appendix S2. Additional notes to the selection of studies and exclusion criteria

Table A1. List of field journals, classified according to their rank in Scimago Journal & Country Rank (SJR).

Table A2. Literature searching criteria adopted in each bibliographic database: example on Google Scholar.

Figure A2. Literature searching criteria.

Figure A3. Trend of empirical researches on trade effects of Non-Tariff Measures (NTMs) over time.

Table A3. List of papers excluded from the meta-analysis.

Appendix S3. Description of the sample

Table A4. Papers included in the empirical analysis.

Table A5. Descriptive statistics for papers included in the sample.

References

Anders, S. and Caswell, J. A. 'Standards-as-barriers versus standards-as-catalysts: Assessing the impact of HACCP implementation on U.S. seafood imports', *American Journal of Agricultural Economics*, Vol. 91, (2009) pp. 310–321.

Arita, S., Beckman, J. and Mitchell, L. 'Reducing transatlantic barriers on US-EU agri-food trade: What are the possible gains?', *Food Policy*, Vol. 68, (2017) pp. 233–247.

Beckman, J. and Arita, S. 'Modeling the interplay between sanitary and phytosanitary measures and tariff-rate quotas under partial trade liberalization', *American Journal of Agricultural Economics*, Vol. 99, (2016) pp. 1078–1095.

Beghin, J. C., Maertens, M. and Swinnen, J. 'Nontariff measures and standards in trade and global value chains', *Annual Review of Resource Economics*, Vol. 7, (2015) pp. 425–450.

Belsley, D. A., Kuh, E. and Welsch, R. E. *Regression Diagnostics* (New York: Wiley, 1980).

Böcker, T. G. and Finger, R. 'A meta-analysis on the elasticity of demand for pesticides', *Journal of Agricultural Economics*, Vol. 68, (2017) pp. 518–533.

Buzby, J. and Mitchell, L. 'Private, national, and international food-safety standards', *Journal of Food Distribution Research*, Vol. 37, (2006) pp. 1.

Card, D. and Krueger, A. B. 'Time-series minimum-wage studies: A meta-analysis', *The American Economic Review*, Vol. 85, (1995) pp. 238–243.

Cardamone, P. 'The effect of preferential trade agreements on monthly fruit exports to the European Union', *European Review of Agricultural Economics*, Vol. 38, (2011) pp. 553–586.

Chen, C., Yang, J. and Findlay, C. 'Measuring the effect of food safety standards on China's agricultural exports', *Review of World Economics*, Vol. 144, (2008) pp. 83–106.

Cioffi, A., Santeramo, F. G. and Vitale, C. D. 'The price stabilization effects of the EU entry price scheme for fruit and vegetables', *Agricultural Economics*, Vol. 42, (2011) pp. 405–418.

Cipollina, M. and Salvatici, L. 'Measuring protection: Mission impossible?', *Journal of Economic Surveys*, Vol. 22, (2008) pp. 577–616.

Cipollina, M. and Salvatici, L. 'Reciprocal trade agreements in gravity models: A meta-analysis', *Review of International Economics*, Vol. 18, (2010) pp. 63–80.

Crivelli, P. and Gröschl, J. 'The impact of sanitary and phytosanitary measures on market entry and trade flows', *The World Economy*, Vol. 39, (2016) pp. 444–473.

Dal Bianco, A., Boatto, V. L., Caracciolo, F. and Santeramo, F. G. 'Tariffs and non-tariff frictions in the world wine trade', *European Review of Agricultural Economics*, Vol. 43, (2016) pp. 31–57.

Dawson, P. J. 'The export–income relationship and trade liberalisation in Bangladesh', *Journal of Policy Modeling*, Vol. 28, (2006) pp. 889–896.

Disdier, A. C. and Head, K. 'The puzzling persistence of the distance effect on bilateral trade', *The Review of Economics and Statistics*, Vol. 90, (2008) pp. 37–48.

Disdier, A. C. and Marette, S. 'The combination of gravity and welfare approaches for evaluating non-tariff measures', *American Journal of Agricultural Economics*, Vol. 92, (2010) pp. 713–726.

Doucouliagos, C. and Stanley, T. D. 'Are all economic facts greatly exaggerated? Theory competition and selectivity', *Journal of Economic Surveys*, Vol. 27, (2013) pp. 316–339.

Fernandes, A. M., Ferro, E. and Wilson, J. S. 'Product standards and firms' export decisions', *The World Bank Economic Review*, (2017) [in press].Author give up to date details if available

Ferro, E., Otsuki, T. and Wilson, J. S. 'The effect of product standards on agricultural exports', *Food Policy*, Vol. 50, (2015) pp. 68–79.

de Frahan, B. H. and Vancauteren, M. 'Harmonisation of food regulations and trade in the Single Market: Evidence from disaggregated data', *European Review of Agricultural Economics*, Vol. 33, (2006) pp. 337–360.

Gourdon, J. *CEPII NTM-MAP: A tool for assessing the economic impact of non-tariff measures*, No. 24, (2014). Author give complete ref details – name and location of publishing instituton, or the name of journal, issue and vol number

Grant, J. H and Arita, S. *Sanitary and Phyto-Sanitary Measures: Assessment, Measurement, and Impact*. IATRC Commissioned Paper No. 21, 2017. Available online at: iatrcweb.org (last accessed 1 February 2018).

Havránek, T. 'Rose effect and the euro: Is the magic gone?', *Review of World Economics*, Vol. 146, (2010) pp. 241–261.

Hoekman, B. and Nicita, A. 'Trade policy, trade costs, and developing country trade', *World Development*, Vol. 39, (2011) pp. 2069–2079.

Hooker, N. H. and Caswell, J. A. 'A framework for evaluating non-tariff barriers to trade related to sanitary and phytosanitary regulation', *Journal of Agricultural Economics*, Vol. 50, (1999) pp. 234–246.

Jayasinghe, S., Beghin, J. C. and Moschini, G. C. 'Determinants of world demand for U.S. corn seeds: The role of trade costs', *American Journal of Agricultural Economics*, Vol. 92, (2010) pp. 999–1010.

Jeppesen, T., List, J. A. and Folmer, H. 'Environmental regulations and new plant location decisions: Evidence from a meta-analysis', *Journal of Regional Science*, Vol. 42, (2002) pp. 19–49.

Jongwanich, J. 'The impact of food safety standards on processed food exports from developing countries', *Food Policy*, Vol. 34, (2009) pp. 447–457.

Li, Y. and Beghin, J. C. 'A meta-analysis of estimates of the impact of technical barriers to trade', *Journal of Policy Modeling*, Vol. 34, (2012) pp. 497–511.

Livingston, M., Osteen, C. and Roberts, D. 'Regulating agricultural imports to keep out foreign pests and disease', *Amber Waves*, Vol. 6, (2008) pp. 16.

Melo, O., Engler, A., Nahuehual, L., Cofre, G. and Barrena, J. 'Do sanitary, phytosanitary, and quality-related standards affect international trade? Evidence from Chilean fruit exports', *World Development*, Vol. 54, (2014) pp. 350–359.

Olper, A. and Raimondi, V. 'Explaining national border effects in the QUAD food trade', *Journal of Agricultural Economics*, Vol. 59, (2008) pp. 436–462.

Otsuki, T., Wilson, J. S. and Sewadeh, M. 'What price precaution? European harmonization of aflatoxin regulations and African groundnut exports', *European Review of Agricultural Economics*, Vol. 28, (2001a) pp. 263–283.

Otsuki, T., Wilson, J. S. and Sewadeh, M. 'Saving two in a billion: Quantifying the trade effect of European food safety standards on African exports', *Food Policy*, Vol. 26, (2001b) pp. 495–514.

Peterson, E., Grant, J. H., Roberts, D. and Karov, V. 'Evaluating the trade restrictiveness of phytosanitary measures on US fresh fruit and vegetable imports', *American Journal of Agricultural Economics*, Vol. 95, (2013) pp. 842–858.

Rose, A. K. and Stanley, T. D. 'A meta-analysis of the effect of common currencies on international trade', *Journal of Economic Surveys*, Vol. 19, (2005) pp. 347–365.

Salvatici, L., Matthews, A. and Scoppola, M. *Trade Impacts of Agricultural Support in the EU*. IATRC Commissioned Paper No. 938-2016-74528, 2017. Available online at: iatrcweb.org/ (last accessed 1 February 2018).

Santeramo, F. G. and Cioffi, A. 'The entry price threshold in EU agriculture: Deterrent or barrier?', *Journal of Policy Modeling*, Vol. 34, (2012) pp. 691–704.

Santeramo, F. G. and Shabnam, N. 'The income-elasticity of calories, macro-and micro-nutrients: What is the literature telling us?', *Food Research International*, Vol. 76, (2015) pp. 932–937.

Schlueter, S. W., Wieck, C. and Heckelei, T. 'Regulatory policies in meat trade: Is there evidence for least trade-distorting sanitary regulations?', *American Journal of Agricultural Economics*, Vol. 91, (2009) pp. 1484–1490.

Sheldon, I. 'Trade and environmental policy: A race to the bottom?', *Journal of Agricultural Economics*, Vol. 57, (2006) pp. 365–392.

Sheldon, I. 'North–South trade and standards: What can general equilibrium analysis tell us?', *World Trade Review*, Vol. 11, (2012) pp. 376–389.

Shepherd, B. and Wilson, N. L. 'Product standards and developing country agricultural exports: The case of the European Union', *Food Policy*, Vol. 42, (2013) pp. 1–10.

Stanley, T. D. 'Wheat from chaff: Meta-analysis as quantitative literature review', *The Journal of Economic Perspectives*, Vol. 15, (2001) pp. 131–150.

Stanley, T. D. 'Beyond publication bias', *Journal of Economic Surveys*, Vol. 19, (2005) pp. 309–345.

Stanley, T. D. and Doucouliagos, H. *Meta-regression Analysis in Economics and Business* (Oxford: Routledge, 2012).

Stanley, T. D. and Doucouliagos, H. 'Meta-regression approximations to reduce publication selection bias', *Research Synthesis Methods*, Vol. 5, (2014) pp. 60–78.

Stanley, T. D. and Jarrell, S. B. 'Meta-regression analysis: A quantitative method of literature surveys', *Journal of Economic Surveys*, Vol. 3, (1989) pp. 161–170.

Stanley, T. D., Doucouliagos, C. and Jarrell, S. B. 'Meta-regression analysis as the socio-economics of economics research', *The Journal of Socio-Economics*, Vol. 37, (2008) pp. 276–292.

Stanley, T. D., Doucouliagos, H., Giles, M., Heckemeyer, J. H., Johnston, R. J., Laroche, P., Nelson, J. P., Paldam, M., Poot, J., Pugh, G., Rosenberger, R. S. and Rost, K. 'Meta-analysis of economics research reporting guidelines', *Journal of Economic Surveys*, Vol. 27, (2013) pp. 390–394.

Swinnen, J. 'Economics and politics of food standards, trade, and development', *Agricultural Economics*, Vol. 47, (2016) pp. 7–19.
Swinnen, J. 'Some dynamic aspects of food standards', *American Journal of Agricultural Economics*, Vol. 99, (2017) pp. 321–338.
Swinnen, J. F. and Vandemoortele, T. 'Trade and the political economy of food standards', *Journal of Agricultural Economics*, Vol. 62, (2011) pp. 259–280.
UNCTAD. *International Classification of Non-Tariff measures, February 2012 version (UNCTAD/DITC/TAB/2012/2)* (New York, Geneva: United Nations, 2012).
United Nations. *World Economic Situation and Prospects* (New York: UN, 2017).
Wilson, J. S. and Otsuki, T. 'Food safety and trade: Winners and losers in a non-harmonized world', *Journal of Economic Integration*, Vol. 18, (2003) pp. 266–287.
Wilson, J. S. and Otsuki, T. 'To spray or not to spray: Pesticides, banana exports, and food safety', *Food Policy*, Vol. 29, (2004) pp. 131–146.
Xavier, C., Gonzales-Barron, U., Paula, V., Estevinho, L. and Cadavez, V. 'Meta-analysis of the incidence of foodborne pathogens in Portuguese meats and their products', *Food Research International*, Vol. 55, (2014) pp. 311–323.
Xiong, B. and Beghin, J. 'Does European aflatoxin regulation hurt groundnut exporters from Africa?', *European Review of Agricultural Economics*, Vol. 39, (2011) pp. 589–609.
Xiong, B. and Beghin, J. 'Disentangling demand-enhancing and trade-cost effects of maximum residue regulations', *Economic Inquiry*, Vol. 52, (2014) pp. 1190–1203.
Yue, C. and Beghin, J. C. 'Tariff equivalent and forgone trade effects of prohibitive technical barriers to trade', *American Journal of Agricultural Economics*, Vol. 91, (2009) pp. 930–941.