

## **The impact of trade and investment agreements on the implementation non-communicable disease policies, 2014-2019: protocol for a statistical study**

Barlow, P.<sup>1</sup> and Allen, L.<sup>2</sup>

1 – Department of Health Policy, London School of Economics and Political Science, London WC2A 2AE; p.barlow@lse.ac.uk; 020 7955 5380.

2 - Department of Clinical Research, London School of Hygiene and Tropical Medicine, London

Keywords: non-communicable disease prevention policy; trade agreements; commercial determinants of health; unhealthy commodities; quasi-experiment

Keywords: trade policy; non-communicable disease prevention; unhealthy commodities; quasi-experiment

Word count: 4,623 words excluding title page, references, figures and tables.

### **Abstract**

#### *Introduction*

Regulating tobacco, alcohol, and unhealthy foods and drinks is a cornerstone of global efforts to combat the Non-Communicable Disease (NCD) pandemic, but implementation of these policies remains slow. It has been suggested that producers of these unhealthy commodities use rules in Trade and Investment Agreements (TIAs) to delay and undermine NCD policy implementation. Yet, there is no systematic empirical evidence linking TIA participation to reduced implementation. Here we present a study protocol for a statistical analysis of the relationship between TIA participation and the implementation of regulations on tobacco, alcohol, and unhealthy food and drink in 154 countries, 2014-2019.

#### *Methods and analysis*

We aim to examine whether participation in TIAs with the EU and US is associated with implementation of regulations targeting tobacco, alcohol, and unhealthy food and drink. We focus on TIAs with these countries as their TIAs create multiple opportunities to contest health regulations, and a majority of the major unhealthy commodity producers are registered in these jurisdictions. Partial and full implementation is captured in a recently published dataset which systematically coded implementation of 11 NCD policies in 2014, 2016 and 2019. We will combine these outcome data with TIA membership and covariate data from multiple sources. We will calculate descriptive statistics and use both regression adjustment and matching to conduct covariate-adjusted, quasi-experimental comparisons of implementation levels and progress according to whether or not countries have a TIA with the EU or US. Further analyses and robustness checks will examine additional TIA participation arrangements and test the sensitivity of our results to our model specifications.

#### *Ethics and dissemination*

Ethics approval will not be required as the study uses anonymised and pre-aggregated data. Findings will be disseminated to policymakers via personal contacts and press releases in parallel with scientific papers and conference presentations.

## **Funding**

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

## Background and study rationale

Non-Communicable Diseases (NCDs) including cardiovascular disease, chronic respiratory diseases, and diabetes kill 41 million people each year, equivalent to 71% of all deaths globally, and are considered a major challenge to sustainable development in the 21<sup>st</sup> century (1,2). To tackle this global health crisis, WHO recommends governments implement a suite of ‘best-buy’ interventions, which are cost-effective means to prevent NCDs, including those summarised in Box 1 (3). Despite unanimous endorsement of these and other interventions targeting unhealthy commodities by all 194 WHO member states, implementation remains slow and uneven (4).

[Box 1 about here]

It has been suggested that the unhealthy commodity industries play a major role in delaying and undermining NCD policy implementation by invoking rules in Trade and Investment Agreements (TIAs) to contest regulations targeting the marketing and promotion of their products (5–9). TIAs include international investment agreements, which are treaties between two or more states for the purpose of promotion and protection of cross-border investments, and bi-lateral or regional free trade agreements, which seek to encourage both trade and investment (10). Between 1960 and 2020, governments worldwide ratified a total of 654 trade agreements and 2,841 investment treaties (11,12).

TIAs empower the unhealthy commodity industries to challenge, delay, and undermine regulations targeting tobacco, alcohol, and unhealthy food and drinks. This can occur as a result of specific commitments that governments make in TIAs, especially those within a new generation of ‘deep’ agreements (13). Deep TIAs expand investor protections and include ‘behind the border’ trade provisions which move beyond a historical focus on border taxes and quotas to harmonize or limit domestic regulations (11). One of the most controversial clauses establishes procedures for a form of international arbitration, referred to as investor-state dispute settlement (ISDS) (14). ISDS and other TIA commitments enable foreign corporations to strategically delay implementation of unhealthy commodity regulations by threatening or initiating a costly legal dispute about a policy (15,16). This delays implementation where governments back down in response to legal threats or disputes, a phenomenon known as ‘regulatory chill’ (15). In addition, deep TIAs can establish avenues for industry input in policy-making and commitments to sharing information about policy proposals (7). This creates opportunities for industry to contest regulations as they are being developed (17).

Delays in regulating unhealthy commodities are especially likely where governments sign TIAs with the US or EU as these countries regularly pursue deep TIAs, and many of the world’s largest and most profitable processed food, tobacco, and alcohol companies are headquartered in these jurisdictions (18,19). Deep TIAs in these jurisdictions bestow domestic companies with a potent means to stall implementation, whilst their substantial profits enable them to mount well-financed legal challenges that are prohibitively costly for many nations to withstand (20,21).

Researchers have identified several instances where multi-national businesses have invoked TIAs to oppose unhealthy commodity regulations. For example, the tobacco giant Philip Morris challenged the tobacco packaging legislation the governments of Australia and Uruguay had adopted by arguing the measures were inconsistent with the countries’ TIAs (8). These formal disputes received significant attention but were ultimately unsuccessful.

However, dispute threats and challenges made outside of courts occur more frequently and can be highly influential, as indicated in analyses of challenges to food, tobacco, soft-drink and alcohol regulations at WTO committee meetings (22). Case study research has identified several instances where these challenges and other dispute threats were followed by decisions to delay implementation (6,23).

Yet, the evidence linking industry pressure via TIA participation to actual policy outcomes remains weak. As far as we are aware, there is no quantitative evidence examining the relationship between TIA participation and policy implementation. Such an analysis requires systematic and harmonized data tracking policy implementation, or a lack thereof, across multiple years and countries. In the absence of such data, research has focussed on specific legal texts and descriptions of trade-related challenges or disputes without systematically connecting these to policy outcomes (17,22,24–27). Qualitative studies have also identified how concerns about legal threats and compliance with TIAs featured in policy-maker decision-making and suspected instances regulatory chill (28). Whilst previous research has been informative, it remains unclear whether TIAs have a causal impact on policy and whether select case studies are generalisable, or are instead high-profile exceptions to a broader tendency for governments to successfully implement policies despite their TIA arrangements and corresponding industry pressure.

These ambiguities are compounded by the recognition within most TIAs that governments have a legitimate need to regulate to protect public health (27,29). Such measures can be therefore deemed permissible according to the agreement, and governments that are aware of such flexibilities or who have a strong political commitment to NCD prevention can sometimes use these protections to withstand business challenges. For example, Chile's successful defence of its front-of-pack labelling legislation against the food industry's claim that the measure was incompatible with WTO rules has been partially attributed to the government's access to legal expertise and its determination to withstand industry pressure (30). Furthermore, businesses operating in a jurisdiction with high levels of regulation may promote increased regulation elsewhere in order to avoid competitive disadvantages associated with higher domestic regulation (31). This possibility became evident when, in 2015, Nestle, Kellogs, Mars, Mondelez supported an EU-wide trans-fat limit in order to create "a common level-playing field for businesses" after similar legislation had been introduced in the US (32). TIAs may therefore promote policy convergence, with the net positive or negative effect dependent on the number and scope of NCD policies in TIA partner countries.

Here we outline a protocol for a statistical analysis of the relationship between TIA participation and the implementation of WHO-recommended regulations on unhealthy commodities. We use repeated cross-national data collected over 3 years (2014, 2016, and 2019) and statistical modelling to examine the relationship between TIA participation under different TIA arrangements, models, and assumptions.

## **Objectives**

The objective of this study is to examine whether participation in TIAs with the US or EU is associated with implementation of policies targeting tobacco, alcohol, and unhealthy foods and drinks. In sensitivity analyses we examine whether these associations apply specifically to US/EU TIAs, where TIAs are often 'deep' in scope and where many large unhealthy commodity producers are registered, or are instead generalizable to a range of other TIA participation arrangements.

## Methods

### *Data and measurement*

Our primary dependent variable of interest captures government implementation of policies targeting the marketing and promotion of tobacco, alcohol and unhealthy food and drink in 2014, 2016 and 2019. Implementation of these regulations is assessed by the WHO using regular NCD country capacity surveys. Cross-sectional implementation survey responses have been published for 194 countries reporting data from 2014, 2016, and 2019. We will use a recently published dataset which systematically coded WHO's policy monitoring data across the three years as fully implemented, partially implemented, not implemented, or no data available. We subset this data to the 11 regulations targeting alcohol, tobacco, and unhealthy food and soft drinks, shown in Box 1.

For each of the 11 measures in Box 1, we create a dichotomous outcome variable indicating whether or not a country has achieved full implementation of the measure, coded as 0 (no implementation) or 1 (full implementation) in the NCD policy dataset. In subsequent analyses we examine whether our results are consistent when using alternative coding procedures (see 'Additional Analyses' below).

### *Independent variable*

Theoretically, implementation of the regulations outlined in Box 1 may be influenced by a range of TIA arrangements. In our primary analysis we examine the relationship between policy implementation and participation in US/EU TIAs, with secondary analyses examining alternative TIA participation indicators, as outlined below (see 'Additional Analyses'). For our primary analysis we create a dichotomous 'treatment' variable indicating whether a country is a member of either a US or EU TIA (1) or not (0). We will create this variable using a list of investment treaties in force, by country-year, from the UN's Investment Policy Hub (12). We complement this with data indicating trade agreements in force, by country-year, from the Design of Trade Agreements Database, DESTA (11).

### *Statistical analyses*

We will first use simple descriptive statistics to examine variation in implementation of unhealthy commodity regulations in countries with and without US/EU TIAs. We then estimate statistical models to adjust for potential confounding in these comparisons.

All statistical analyses will be performed in R. Statistical significance will be assessed based on whether the p-value for each test is less than 0.05 ( $p < 0.05$ ). We will also report 95% confidence intervals.

### *Descriptive analysis*

Two-proportion z-tests will be used to assess differences in the proportion of countries that have achieved full implementation of each of the unhealthy commodity regulations in Box 1 according to whether or not they have TIAs with the US or EU. Z-tests will also be performed to compare the mean number of fully implemented regulations in each commodity category in countries with and without US/EU TIAs.

### *Specifications to minimise confounding*

Countries which do and do not participate in US/EU TIAs differ from one-another in ways that are associated with NCD policy implementation, for example with respect to democratization and GDP per capita (33,34). These differences may account for variation in

policy implementation across countries with and without US/EU TIAs, as captured in the descriptive statistics above. To address this we adopt a quasi-experimental approach, which is appropriate for testing descriptive causal hypotheses when randomization is unfeasible, as is the case with TIAs (35).

There are multiple different ways to approach our analysis, and in order to obtain the most robust findings, we intend to employ three widely used statistical approaches that are feasible with our data, report diagnostic tests, and base on our main results on best performing models and most consistent results. The three approaches are:

- 1) regression adjustment,
- 2) matching, and
- 3) first-difference models.

Each approach has different advantages, disadvantages and data requirements (36–38). We will evaluate the performance of each model, and the main results presented in the paper will be those that are based on the best performing models with respect to statistical power and additional diagnostic tests, as outlined in further detail below. Data, code, and results from all models will be reported in supplementary appendices in order to mitigate against selective reporting.

#### Step 1: regression adjustment

Our baseline logistic regression model is:

$$\text{Equation 1.} \quad Y_{it} = B_0 + B_1TIA_{it-1} + B_2X_{it-1} + B_3Wave_t + e_{it}$$

Where  $Y_{it}$  is the binary indicator of policy implementation of a given regulation in country  $i$  in year  $t$  (2014, 2016 or 2019), and  $B_0$  is the intercept.  $TIA_{it}$  is the indicator of US/EU TIA participation in country  $i$  in year  $t-1$  with coefficient  $B_1$ . To allow for a delayed effect, we lag this indicator by one year.  $X_{it-1}$  in Equation 1 is a vector of controls measured in year  $t-1$  with coefficients in the vector  $B_2$ . We control for variables known to be related to both treatment assignment and the outcome, and which are not affected by the treatment (39,40). We incorporate the following relevant and measurable covariates: democratization, GDP per capita, the share of the population of secondary education age that is enrolled in secondary education, implementation of non-trade business regulations, WTO participation, geographic region, participation in TIAs where other major unhealthy commodity producers are headquartered, and international political integration (or ‘political globalization’). Again these are lagged by 1 year in our models to allow for a delayed effect. Table 1 lists the data sources and measurement of these variables, and the rationale for their inclusion according to studies examining determinants of US/EU TIA ratification and NCD policy implementation (11,22,47,33,34,41–46).

[Table 1 about here]

$Wave_t$  in Equation 1 is a control for the wave of data collection with coefficient  $B_3$ . This adjusts for unobserved macro-economic and political factors which influence policy implementation, vary across data collection waves, and are common across all countries.  $e_{it}$  in Equation 1 is the error term. We estimate Equation 1 using pooled logistic regression models and account for within-country correlations in the error term by clustering the standard errors at the country level.

Finally, we use the estimated coefficients in this model to calculate predicted probabilities of full implementation of each regulation in countries with and without US/EU TIAs as well as the Average Marginal Effect of US/EU TIAs, that is, the difference in these predicted probabilities (48).

### Step 2: matching

Regression adjustment does not fully correct for imbalance in covariates across countries with different TIA arrangements and can lead to inferences that are beyond the bounds of what is observed within the data (49,50). Matching has been widely applied to address these issues in analyses of the economic, environmental, and health consequences of TIAs (33,51–54). Matching pre-processes the data to identify the ‘untreated’ comparison country or countries that is/are most similar to each country with a EU/ US TIA, thereby reducing imbalance in covariates (40). After the matching, the quality of matches and post-matching covariate balances can be analysed using a range of diagnostic tests. The matched sample is then analysed using regression models with controls for any remaining covariate imbalances (55).

There are many different matching algorithms, and we plan to employ the multiple procedures that are feasible with our data, report diagnostic tests, and base on our main results on best performing models and most consistent results (40). When first assessing ‘similarity’ between countries to identify relevant matches, we begin with one of the most commonly used, unit-free metrics: the propensity score. We estimate a logit model to predict treatment assignment (ie US/EU TIA participation) based on the set of covariates used to estimate Equation 1. Predicted probabilities of US/EU TIA participation are then calculated from this model; these are the ‘propensity scores’. We then use three alternatives to traditional propensity score matching: i) the non-parametric covariate-balancing propensity score algorithm, which identifies propensity scores such that both covariate balance and prediction of treatment assignment are maximized, and ii) calculation of an alternative measure of similarity/ dissimilarity, the Mahalanobis distance.

Next, for each procedure, we select comparison countries based on whether their propensity scores are similar to the ‘treated’ countries with a US/EU TIAs. We use ‘nearest-neighbour’ matching to identify the 1- and 3-, and 5- most similar untreated countries. Since nearest matches on propensity scores may nevertheless be dissimilar to the countries with US/EU TIAs, we will set a caliper defining the maximum difference in propensity score to make a country match (‘radius matching’); in our case, we will use a caliper of 0.1 (i.e., ten percentage points of the likelihood of becoming a US/EU TIA member). This choice is necessarily arbitrary, and we also use alternative caliper specifications and choose these where they yield better performance on our diagnostic tests (see below).

As stated above, diagnostic tests will be used to determine which of the above models performs best and hence is to be used to estimate US/EU TIA impacts on policy. Following recommendations in Stuart (2010) and King and Zeng (2006), these tests include examination of the number of successfully matched pairs, covariate balance tests, comparing the difference in covariate means and variance, and tests of common support which examine the distribution of propensity scores across treated/ untreated units and the ‘convex hull’ of the covariates (40,50). We will include all diagnostic test results in the study write-up and document the rationale for our final selection.

Finally, as a robustness check we will weight observations in inverse proportion to their similarity in the propensity score (as calculated from the best performing model above), rather than select a limited number of comparison countries, so that more ‘untreated’ units are used but more dissimilar units are given lower weight (‘kernel matching’). This approach increases statistical power, which can be limited with the aforementioned approaches, although it comes at a cost of incorporating some comparison countries which may be dissimilar to the countries with EU/US TIAs.

### Step 3: first difference models

The models described above control for measurable confounding/ covariate imbalance but do not address the influence of unobserved or unmeasured factors that predict both US/EU TIA participation and NCD policy implementation. A common strategy to reduce the influence of time-invariant unobserved/unmeasurable confounders is to estimate fixed-effects models (38). However, with only 3 years of data we lack sufficient data to estimate sufficiently powered fixed-effects models. We nevertheless exploit the temporal nature of our data by modelling changes over time in the outcome between two data collection waves (2015-2017 and 2017-2020). According to An and Winship (2015), this strategy is advantageous when few years of data are available as first differencing the outcomes helps to remove the influence of unobserved time-invariant factors, whilst posing fewer data restrictions than fixed-effects models. This approach is therefore recommended by An and Winship and has previously been applied to study political and macro-economic influences on health with few years of data (56–58).

After first-differencing the outcome, we apply the matching and weighting procedures outlined above on the differenced outcomes, modelling changes in implementation between two periods as a function of US/EU TIA participation and covariates in the first period.

### *Patient and Public Involvement*

No patient involved.

### **Additional analyses**

We conduct additional analyses to examine heterogeneity in the relationship between TIA participation and unhealthy commodity regulations across countries. We further conduct a number of tests which evaluate i) whether our results are consistent using alternative outcome and TIA participation indicators, ii) the specificity of our results, and iii) whether our results are robust when using alternative model specifications.

### *Heterogeneity analyses*

Our heterogeneity analyses examine whether the relationships examined above vary depending on the income level of the country and state capacity (indicators in Table 2 below), both of which have previously been identified as important modifiers of the relationship between TIAs and domestic health policy (20,59). We also examine if our results pertain specifically to US TIAs, since US companies have regularly been cited as the main actors that used strategic litigation to oppose NCD policies (8,26,60). For comparison we will conduct a separate analysis to assess if our results pertain specifically to EU TIAs.

### *Alternative TIA participation measures*

As noted above, implementation of unhealthy commodity regulations may be influenced by a range of TIA arrangements and the relationships we examine above may not apply



exclusively to US/EU TIAs. We therefore re-estimate the relationship between TIA participation and unhealthy commodity regulations using several alternative TIA participation indicators. As with our main models we use regression adjustment, matching, and first-difference models to perform these tests. However, some additional analyses examine continuous rather than binary indicators of TIA participation (e.g. number of TIAs rather than participation in either a US or EU TIA). Since matching is not appropriate for analyses of continuous treatment indicators, we use an alternative but similar approach to address covariate imbalance: non-parametric Covariate Balancing Generalized Propensity Score (npCBGPS) estimation (61). npCBGPS uses an algorithm to search for a set of country-weights which, when applied to the data, minimises the correlation between the covariates and the probability of a given number of TIAs, whilst simultaneously maximising treatment prediction. We then apply these weights to the data when estimating our regression models. We estimate ‘doubly robust’ regression models, which incorporate the weights as well as the regression controls specified above (62).

We first re-estimate our models replacing the indicator of participation in a US/EU TIA with the following:

- i. The *total number* of TIAs with the US and EU that a country participates in,
- ii. The number of TIAs *with the US only* that a country participates in,
- iii. The number of TIAs *with the EU only* that a country participates in.

Second, we examine whether our results apply to all TIAs with countries that are major producers unhealthy commodities, rather than the US/EU specifically. To assess this we will create variables which capture whether or not a country has a TIA with at least one other country where a major producer of unhealthy commodities is registered, with a second variable indicating the number of these TIAs. To identify these countries, we will use data from the 2014, 2016 and 2019 *Forbes* Global 2000 reports showing where the 10 largest, publicly listed companies in the alcohol, tobacco, and food/ soft-drink sectors were registered in those years (63). We will create three separate ‘treatment’ (TIA participation) variables when examining implementation of policies targeting each category of commodities: i) tobacco, ii) alcohol, and iii) unhealthy food & drink. For each commodity, we code TIA participation as ‘1’ if a country has at least one TIA with a country in which one of these largest top 10 producers of the commodity in question is headquartered, or 0 otherwise. In additional analyses we will count the number of TIAs with these countries to assess how this is associated with unhealthy commodity regulation.

Third, we examine the total number of TIAs a country participates in, which creates exposure to multi-national companies operating in multiple jurisdictions. In addition, having multiple TIAs creates exposure to ‘venue’ or ‘forum shopping’, whereupon businesses use the TIA which is most advantageous to their case (6,64,65). We assess this by calculating the total number of TIAs in each country in each year (2014, 2016 and 2019). In subsequent analyses we interact the number of TIAs with the number of TIA partners to examine variation according to the number of distinct TIA partners, as several countries have multiple agreements with one partner country.

Finally, there is likely to be a strong correlation between each of the TIA participation indicators described above (for example participation in US TIAs is associated with the total number of TIAs), and so our analysis of one indicator may capture the influence of another. Furthermore, the association between each indicator and policy implementation may vary in magnitude. We therefore estimate a final model in which we incorporate the following TIA

participation indicators simultaneously in order to adjust for this co-variation and compare the magnitude of their estimated impacts: total number of TIAs, US TIA participation, EU TIA participation, and participation in a TIA with a major producer of the unhealthy commodity in question. This model also adjusts for the same covariates as our baseline model, as described above.

#### *Alternative outcome measures*

We examine several alternative outcome indicators. As noted above, TIAs could promote policy convergence, leading to more similar regulations among TIA partners, and whether this leads to a lower or higher level of regulation may depend on the specific regulations in force in TIA partner countries. In a first set of analyses we re-estimate our models substituting the outcome variable with an indicator of similarity in implementation in a given country compared to the mean in the TIA partner countries (the EU, US, or across all TIA partners, depending on the TIA participation indicator being evaluated). This indicator of similarity in implementation is calculated by dividing the implementation score in TIA partner countries with the implementation score in a given country with (or without) a TIA, so that values closer to 1 indicate more similar implementation levels.

We further re-estimate the above models using an alternative indicator capturing either no implementation (0), or at least partial implementation, i.e., coded as partial or full implementation in the NCD policy dataset (1), rather than full implementation only. We will further count the number of partially/ fully implemented regulations across all commodities and each category of commodities, and examine the association between participation in US/EU TIAs and the number of regulations partially/ fully implemented using a Poisson regression model.

Finally, we will create variables capturing whether countries move from no to partial/full implementation, or from partial to full implementation, between 2014 and 2016 and 2016 and 2019, and examine the association between participation in US/EU TIAs and each of these variables.

#### *Sensitivity analyses*

We first conduct analyses to assess whether our results are sensitive to our propensity score model specification and our decision to use this metric rather than alternative measures of similarity between countries (40,66). First, Stuart (2010) recommends choosing a parsimonious list of variables when estimating propensity score models in small samples like ours, as including a large number of variables can lead to increases in variance. However, the results from our matching models may be sensitive to our choice of covariates and limited selection thereof. We therefore conduct a test in which we incorporate additional variables that predict US or EU TIA participation: distance, sharing a common border, sharing a common language. We also re-estimate all models (regression and matching) incorporating a further possible covariates of TIA formation and NCD policy to assess the sensitivity of all models to the inclusion of these additional variables: the total number of TIAs in force in a country, and social globalization (67).

Second, matching on the propensity score has the advantage of using a single indicator to perform matches, rather than multiple variables across which matches may not be possible, a phenomenon known as the ‘curse of dimensionality’. However, a disadvantage of propensity score matching is that covariate imbalances may not be sufficiently reduced (40,66). We

therefore conduct an additional analysis in which we use an alternative metric recommended by King and Nielsen: the Mahalanobis distance. This metric is calculated by taking a weighted-measure of similarity in covariates in treated and untreated units.

Next, we assess whether our results may be attributable to alternative explanations. We re-estimate our models with a control for the number of fully implemented measures in the NCD policy dataset that did not directly regulate unhealthy commodities. This control captures other influences on implementation which may bias our estimates. We further conduct a placebo analysis, in which we re-estimate our models examining implementation of an NCD policy that we would not expect to be affected by TIAs: whether or not a country has set time-bound national targets to address NCDs based on WHO guidance. This helps to identify whether our results are driven by TIA implementation or some third factor leading to general changes in implementation.

Table 2 below lists the data sources and measurement of the variables included in our additional analyses.

[Table 2 about here]

### *Missing data*

For all analyses outlined above we will apply listwise deletion where data are missing. However, this approach can produce biased results where data are missing not at random (68). In addition, listwise deletion markedly reduces the sample size, which in turn causes a substantial loss of precision and power. Finally, the developers of the NCD policy implementation dataset assigned a score of zero for those implementation policies that are not reported on the basis that if a country is unable to determine whether a policy has been implemented, then it is likely that the policy was not implemented. However, it remains possible that assigning a score of zero in such cases will exert a downward bias on our estimates (for example if countries successfully implemented policies and were coded as zero in the data).

To address these limitations we will re-estimate our models using imputed outcome and covariate data. To impute these missing data we will use a multivariate normal model (with log transformations for non-normal variables) incorporating all covariates in our original model. The number of imputations will follow von Hippel's rule of thumb: that the number of imputations ( $m$ ) should be similar to the percentage of cases that are incomplete. Hence  $m$  will be determined based on the number of missing cases in the data (69). We will then re-estimate our models using the imputed datasets and combine the results from these models, using Rubin's Rules to calculate the Standard Errors.

### **Ethics and dissemination**

Ethical approval is not necessary as all data are publicly available. Furthermore, all data are collected at the country-level and it is not possible to identify individuals in these aggregated data. Findings will be disseminated via scientific papers and conference presentations.

### **Conclusion**

Using a novel dataset capturing NCD policy implementation and a range of statistical models, this study will provide new insights into the relationship between TIA participation and implementation of regulations targeting unhealthy commodities. The findings will help researchers and policymakers better understand whether TIAs actually constrain implementation of regulations targeting unhealthy commodities, and hence whether TIAs may need to be modified in order to facilitate government efforts to curb the global non-communicable disease pandemic.

## References

1. UN. Political Declaration of the Third High-level Meeting of the General Assembly on the Prevention and Control of Non-communicable Diseases. A/73/L.2. Geneva, Switzerland: United Nations; 2018.
2. WHO. Fact sheet on noncommunicable diseases [Internet]. 2021 [cited 2021 Jul 15]. Available from: <https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases>
3. WHO. Tackling NCDs: ‘Best buys’ and other recommended interventions for the prevention and control of noncommunicable diseases. Geneva, Switzerland: WHO; 2017.
4. WHO. 2021 Progress report on the Global Action Plan for Healthy Lives and Well-being for All. In Geneva, Switzerland: WHO; 2021.
5. Hawkins B, Holden C. A Corporate Veto on Health Policy? Global Constitutionalism and Investor-State Dispute Settlement. *J Health Polit Policy Law*. 2016/06/02. 2016 Oct;41(5):969–95.
6. Crosbie E, Glantz SA. Tobacco industry argues domestic trademark laws and international treaties preclude cigarette health warning labels, despite consistent legal advice that the argument is invalid. *Tob Control*. 2014 May 1;23(3):e7 LP-e7.
7. Friel S, Hattersley L, Townsend R. Trade Policy and Public Health. *Annu Rev Public Health*. 2015 Jan 15;36(1):325–44.
8. Jarman H. Attack on Australia: Tobacco industry challenges to plain packaging. *J Public Health Policy*. 2013 Mar 21;34(3):375–87.
9. Barlow P, Thow A. Neoliberal discourse, actor power, and the politics of nutrition policy: a qualitative analysis of informal challenges to nutrition labelling regulations at the World Trade Organization, 2007-2019. *Soc Sci Med*. 2021;
10. McNeill D, Barlow P, Birkbeck CD, Fukuda-Parr S, Grover A, Schrecker T, et al. Trade and investment agreements: Implications for health protection. *J World Trade*. 2017;51(1).
11. Dür A, Baccini L, Elsig M. The design of international trade agreements: Introducing a new dataset. *Rev Int Organ*. 2014;9(3):353–75.
12. UNCTAD. Investment Policy Hub: Investment Agreements Navigator [Internet]. 2021 [cited 2021 Jul 7]. Available from: <https://investmentpolicy.unctad.org/international-investment-agreements>
13. Baldwin R. 21st Century Regionalism: Filling the gap between 21st century trade and

- 20th century trade rules. [Geneva, Switzerland: World Trade Organisation]: Staff Working Paper ERSD-2011-08; 2011.
14. Bernasconi-Osterwalder N, Cosbey A, Johnson L. Investment treaties and why they matter to sustainable development: Questions and answers. International Institute for Sustainable Development; 2000.
  15. Tienhaara K. Regulatory chill and the threat of arbitration: A view from political science. In: Brown C, Miles K, editors. *Evolution in Investment Treaty Law and Arbitration*. Cambridge: Cambridge University Press; 2011. p. 606–28.
  16. Schram A, Ruckert A, VanDuzer JA, Friel S, Gleeson D, Thow A-M, et al. A conceptual framework for investigating the impacts of international trade and investment agreements on noncommunicable disease risk factors. *Health Policy Plan*. 2017;33(1):123–36.
  17. Thow A-M, Snowdon W, Labonté R, Gleeson D, Stuckler D, Hattersley L, et al. Will the next generation of preferential trade and investment agreements undermine prevention of noncommunicable diseases? A prospective policy analysis of the Trans Pacific Partnership Agreement. *Health Policy (New York)*. 2015;119(1):88–96.
  18. Horn H, Mavroidis PC, Sapir A. Beyond the WTO? An Anatomy of EU and US Preferential Trade Agreements. *World Econ*. 2010;33(11):1565–88.
  19. Allen LN, Hatefi A, Feigl AB. Corporate profits versus spending on non-communicable disease prevention: an unhealthy balance. *Lancet Glob Heal*. 2019 Nov 1;7(11):e1482–3.
  20. McNeill D, Birkbeck C, Fukuda-Parr S, Grover A, Schrecker T, Stuckler D. Political origins of health inequities: trade and investment agreements. *Lancet*. 2017;389(10070):760–2.
  21. Stuckler D, McKee M, Ebrahim S, Basu S. Manufacturing epidemics: The role of global producers in increased consumption of unhealthy commodities including processed foods, alcohol, and tobacco. *PLoS Med*. 2012;9(6):10.
  22. Barlow P, Labonte R, McKee M, Stuckler D. Trade challenges at the World Trade Organization to national noncommunicable disease prevention policies: A thematic document analysis of trade and health policy space. *PLoS Med*. 2018;15(6):e1002590.
  23. Allen LN. Commercial Determinants of Global Health. In: *Handbook of Global Health*. Springer International Publishing; 2020. p. 1–37.
  24. O'Brien P, Mitchell AD. On the bottle: health information, alcohol labelling and the WTO Technical Barriers to Trade Agreement. *QUT L Rev*. 2018;18:124.
  25. Shaffer ER, Brenner JE, Houston TP. International trade agreements: a threat to tobacco control policy. *Tob Control*. 2005;14(suppl 2):ii19–25.
  26. Crosbie E, Carriedo A, Schmidt L. Hollow Threats: Transnational Food and Beverage Companies' Use of International Agreements to Fight Front-of-Pack Nutrition Labeling in Mexico and Beyond. *Int J Heal Policy Manag*. 2020;
  27. Mitchell A, Voon T. Implications of the World Trade Organization in combating non-communicable diseases. *Public Health*. 2011;125(12):832–9.
  28. Schram A, Friel S, Anthony VanDuzer J, Ruckert A, Labonté R. Internalisation of International Investment Agreements in Public Policymaking: Developing a Conceptual Framework of Regulatory Chill. *Glob Policy*. 2018 May 1;9(2):193–202.

29. McGrady B. Trade and public health: the WTO, tobacco, alcohol, and diet. Cambridge University Press; 2011.
30. Dorlach T, Mertenskötter P. Interpreters of International Economic Law: Corporations and Bureaucrats in Contest over Chile's Nutrition Label. *Law Soc Rev.* 2020 Sep 1;54(3):571–606.
31. Parziale A, Ooms G. The global fight against trans-fat: the potential role of international trade and law. *Global Health.* 2019;15(1):46.
32. BEUC. Food giants join BEUC call to limit harmful trans-fats [Internet]. 2015 [cited 2021 Jul 15]. Available from: <https://www.beuc.eu/press-media/news-events/food-giants-join-beuc-call-limit-harmful-trans-fats>
33. Baier SL, Bergstrand JH. Estimating the effects of free trade agreements on international trade flows using matching econometrics. *J Int Econ.* 2009;77(1):63–76.
34. Allen LN, Nicholson BD, Yeung BYT, Goiana-da-Silva F. Implementation of non-communicable disease policies: a geopolitical analysis of 151 countries. *Lancet Glob Heal.* 2020 Jan 1;8(1):e50–8.
35. Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Heal.* 2012;66(12):1182–6.
36. Athey S, Imbens GW. The state of applied econometrics: Causality and policy evaluation. *J Econ Perspect.* 2017;31(2):3–32.
37. Abadie A, Cattaneo MD. *Econometric Methods for Program Evaluation.* *Annu Rev Econom.* 2018 Mar 13;10(1):465–503.
38. Angrist J, Pischke J-S. *Mostly Harmless Econometrics: An Empiricist's Companion.* Oxford, UK: Princeton University Press; 2009.
39. Morgan SL, Winship CCN-H. *M 2007 300. 7. 22. Counterfactuals and causal inference: methods and principles for social research.* Cambridge: Cambridge University Press; 2007.
40. Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci.* 2010 Feb 1;25(1):1–21.
41. Egger H, Egger P, Greenaway D. The trade structure effects of endogenous regional trade agreements. *J Int Econ.* 2008;74(2):278–98.
42. Milner H V, Kubota K. Why the Move to Free Trade? Democracy and Trade Policy in the Developing Countries. *Int Organ.* 2005/02/15. 2005;59(1):107–43.
43. Furtado KS, Budd EL, Ying X, deRuyter AJ, Armstrong RL, Pettman TL, et al. Exploring political influences on evidence-based non-communicable disease prevention across four countries. *Health Educ Res.* 2018 Apr 1;33(2):89–103.
44. Thow AM, Abdool Karim S, Mukanu MM, Ahaibwe G, Wanjohi M, Gaogane L, et al. The political economy of sugar-sweetened beverage taxation: an analysis from seven countries in sub-Saharan Africa. *Glob Health Action.* 2021;14(1):1909267.
45. Dobbin F, Simmons B, Garrett G. The global diffusion of public policies: Social construction, coercion, competition, or learning? *Annu Rev Sociol.* 2007;33:449–72.
46. Valente TW, Pitts S, Wipfli H, Vega Yon GG. Network influences on policy implementation: Evidence from a global health treaty. *Soc Sci Med.* 2019;222:188–97.

47. Busch ML. Overlapping Institutions, Forum Shopping, and Dispute Settlement in International Trade. *Int Organ*. 2007 Feb 13;61(4):735–61.
48. Leeper TJ. Interpreting regression results using average marginal effects with R's margins. Available *Compr R Arch Netw*. 2017;
49. Abadie A, Cattaneo MD. Econometric Methods for Program Evaluation. *Annu Rev Econom*. 2018 Mar;10(1):465–503.
50. King G, Zeng L. The dangers of extreme counterfactuals. *Polit Anal*. 2006;14(2):131–59.
51. Cowling K, Stuart EA, Neff RA, Vernick J, Magraw D, Pollack Porter K. The relationship between joining a US free trade agreement and processed food sales, 2002–2016: a comparative interrupted time-series analysis. *Public Health Nutr*. 2020/03/19. 2020;23(9):1609–17.
52. Baghdadi L, Martinez-Zarzoso I, Zitouna H. Are RTA agreements with environmental provisions reducing emissions? *J Int Econ*. 2013;90(2):378–90.
53. Egger P, Larch M, Staub KE, Winkelmann R. The trade effects of endogenous preferential trade agreements. *Am Econ J Econ Policy*. 2011;3(3):113–43.
54. Spilker G, Böhmelt T. The impact of preferential trade agreements on governmental repression revisited. *Rev Int Organ*. 2013;8(3):343–61.
55. Ho DE, Imai K, King G, Stuart EA. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit Anal*. 2007;15(3):199–236.
56. An W, Winship C. Causal Inference in Panel Data With Application to Estimating Race-of-Interviewer Effects in the General Social Survey. *Sociol Methods Res*. 2015 Aug 26;46(1):68–102.
57. Reeves A, Karanikolos M, Mackenbach J, McKee M, Stuckler D. Do employment protection policies reduce the relative disadvantage in the labour market experienced by unhealthy people? A natural experiment created by the Great Recession in Europe. *Soc Sci Med*. 2014;121:98–108.
58. Barlow P, Reeves A, McKee M, Stuckler D. Employment relations and dismissal regulations: Does employment legislation protect the health of workers? *Soc Policy Adm*. 2019 Feb 25;53(939–957).
59. Barlow P. Global disparities in health-systems financing: A cross-national analysis of the impact of tariff reductions and state capacity on public health expenditure in 65 low- and middle-income countries, 1996–2015. *Heal Place*. 2020;63:102329.
60. Crosbie E, Thomson G. Regulatory chills: tobacco industry legal threats and the politics of tobacco standardised packaging in New Zealand. *N Z Med J*. 2018;131(1473):25.
61. Fong C, Hazlett C, Imai K. Covariate balancing propensity score for a continuous treatment: application to the efficacy of political advertisements. *Ann Appl Stat*. 2018;12(1):156–77.
62. Bang H, Robins JM. Doubly Robust Estimation in Missing Data and Causal Inference Models. *Biometrics*. 2005 Dec 1;61(4):962–73.
63. Forbes. Forbes Global 2000 (multiple years) [Internet]. 2021 [cited 2021 May 27].

Available from: <https://www.forbes.com/lists/global2000/#61d7b6a45ac0>

64. Eckhardt J, Holden C, Callard CD. Tobacco control and the World Trade Organization: mapping member states' positions after the framework convention on tobacco control. *Tob Control*. 2016 Nov 1;25(6):692 LP – 698.
65. Eckhardt J, de Biere D. Boomerangs over lac Léman: transnational lobbying and foreign venue shopping in WTO dispute settlement. *World Trade Rev*. 2015;14:507.
66. King G, Nielsen R. Why propensity scores should not be used for matching. *Polit Anal*. 2019;27(4):435–54.
67. Gygli S, Haelg F, Potrafke N, Sturm J-E. The KOF Globalisation Index – revisited. *Rev Int Organ*. 2019 Mar 13;
68. Little RJA, Rubin DB. *Statistical analysis with missing data*, 3rd edition. John Wiley & Sons; 2019.
69. Von Hippel PT. 8. How to impute interactions, squares, and other transformed variables. *Sociol Methodol*. 2009;39(1):265–91.



## Boxes and Tables

*Box 1. WHO-recommended ‘best-buy’ policies targeting the marketing, composition, and consumption of unhealthy commodities*

### Tobacco

- Tobacco taxes
- Smoke-free place policies
- Graphic warnings on cigarette packages
- Tobacco advertising bans

### Alcohol

- Alcohol sales or advertising restrictions
- Alcohol taxes

### Unhealthy food and non-alcoholic beverages

- Legislation implementing the International Code of Marketing of Breast-milk Substitutes.
- Policies to reduce salt/ sodium consumption
- Policies to limit saturated fatty acids and eliminate trans-fats
- Policies targeting the marketing of foods and non-alcoholic beverages to children

*Table 1. Covariates and measurement*

<b>Covariate</b>	<b>Rationale for inclusion</b>	<b>Measure, data type, range, and source</b>	<b>Data availability<sup>a</sup></b>
Democratization	<p>Democracies tend to have more liberal trade policies, and pairs of democracies are more likely to form TIAs. Democratization is also associated with higher NCD implementation scores.</p> <p>Source: Mansfield and Milner (2012); Allen et al. (2020)</p>	<p>Varieties of Democracy (V-Dem) Multiplicative polyarchy index, which combines scores across multiple indicators of democracy (suffrage, free and fair elections, elected officials, freedom of civil and political organization, and freedom of expression).</p> <p>Continuous (0-1), from V-Dem Institute, University of Gothenberg</p>	192-194 countries
GDP per capita	<p>Countries with larger levels of GDP per capita tend to be more open, and countries with similar levels of GDP per capita and with are more likely to ratify TIAs. Countries with higher levels of GDP per capita also have more resources to implement NCD policies.</p> <p>Sources: Egger, Egger &amp; Greenaway (2008); Baier and Bergstrand (2009), Allen et al. (2020)</p>	<p>GDP per capita, PPP (constant 2017 international \$).</p> <p>Continuous (731 to 155,201), from World Bank World Development Indicators</p>	178-194 countries
Human capital	<p>Differences in skilled labour affect whether two countries ratify TIAs, whilst the school enrolment rate is a dimension of human capital that is associated with NCD policy implementation</p> <p>Sources: Egger, Egger &amp; Greenaway (2008); Allen (2020)</p>	<p>The ratio of total enrolment in secondary education (among persons of any age) to the population of the age group that is typically enrolled in secondary education.</p> <p>Continuous (11.01 to 164), from World Bank World Development Indicators</p>	122-141 countries
Government stance on free market capitalism and liberalisation	<p>Support for market liberalization and free markets is associated with trade liberalisation, whilst qualitative/ case-study research indicates that belief in the superiority of individual enterprise and preference against state intervention reduces political willingness to regulate unhealthy commodities</p>	<p>Orientation of the chief executive/ their party with respect to economic policy,</p> <p>Nominal (Left, Right, or Centrist), from the Database of Political Institutions, 2020 edition <a href="https://mydata.iadb.org/">https://mydata.iadb.org/</a></p>	126-131 countries

	Sources: Dutt and Mitra (2002), Milner and Judkins (2004), Thow et al. (2021), Barlow and Thow (2021)		
Business regulations and domestic market liberalisation	Trade liberalisation is associated with other reforms to liberalise domestic markets, including limited business regulation, whilst support for minimal business regulation and domestic market liberalisation associates with opposition to, and limited implementation of, unhealthy commodity regulations  Sources: Billmeier and Nannicini (2013), Rodriguez and Rodrik (2001), Cullerton et al. (2016)	The extent to which the regulatory and infrastructure environments constrain the efficient operation of businesses (excluding trade policies), based on a weighted sum of 13 indicators in the World Bank's Doing Business report.  Continuous (0-100), from The Heritage Foundation	186 countries
WTO Participation	WTO participation may incentivise RTA negotiation as governments seek to further liberalise trade (beyond what is possible within WTO), whilst WTO membership may independently influence NCD prevention  Sources: Baldwin 2009, Barlow et al. 2018	WTO membership status Binary, from Head et al. (2020) and accessed via Centre d'Etudes Prospectives et d'Informations  Binary (0 or 1)	194 countries
Participation in TIAs with other unhealthy commodity exporters (Switzerland, Australia, Canada, New Zealand)	Participation in US/EU TIAs associates with TIA ratification with other unhealthy commodity exporters (as, for example, the US/EU may want to secure the same market advantages as other states), whilst other unhealthy commodity exporters have been found to frequently challenge regulations targeting unhealthy foods, soft-drinks, tobacco and alcohol.  Sources: Barlow et al. 2018; Dur & Baccini 2014	Binary indicator of participation in a TIA with Switzerland, Australia, Canada or New Zealand, as identified Bilateral Investment Treaties listed in the UN's Investment Policy Hub and FTAs in the Design of Trade Agreements Database, DESTA.  Binary (0 or 1)	194 countries
International political integration	International political integration associates with exchange of information, norms and political pressure regarding both trade policy and NCD prevention,	KOF Political Globalisation index, calculated using the number of foreign embassies in a country, personnel contributed to UN security	194 countries

	<p>potentially leading governments to seek and negotiate TIAs whilst also influencing policies targeting unhealthy commodities.</p> <p>Sources: Dobbin et al. 2007; Valente et al. 2019</p>	<p>council missions (%), the number of internationally oriented nongovernmental organisations (NGO) operating in that country, the number of multilateral treaties signed since 1945, the number of memberships in international organizations and a measure for the treaty partner diversity.</p> <p>Continuous, 0-100. From KOF Swiss Economic Institute</p>	
International economic integration	<p>Countries with US/EU TIAs also tend to have larger trade and investment volumes (independent of the effects of US/EU TIAs) as, for example, the importance of trade to the national economy can lead countries to pursue TIAs, and countries may limit/ be pressured to limit policies targeting unhealthy commodities in an attempt to minimize trade disruptions.</p> <p>Source: Baier &amp; Bergstrand (2009); Barlow et al. 2018</p>	<p>KOF de facto Economic Globalisation index, calculated using total trade in goods and services, trade diversity, FDI and portfolio investment, debt, reserves, and international income payments.</p> <p>Continuous, 0-100. From KOF Swiss Economic Institute.</p>	194 countries
Number of TIAs with other countries	<p>Countries with US/ EU TIAs may also tend to have TIAs with other countries, reflecting a general support for trade and market-oriented policies, and these other TIAs may also contain clauses that businesses can make use of to pressure countries to change policies (so-called ‘forum shopping’), as their now global supply chains means they may have operations in many countries.</p> <p>Source: Busch (2007)</p>	<p>DESTA and UNCTAD (as per main variables)</p> <p>Continuous</p>	194 countries

Notes: a: all variables are lagged by one year, with the exception of secondary education for which we lag the variable by 2 years in 2020 due to limited data availability in 2019. Table shows data availability 2014, 2016 and 2019. Ranges show min to max in this period if not equal across years

*Table 2. Data sources for additional analyses*

<b>Covariate</b>	<b>Measure, data type, and data source</b>	<b>Data availability</b>
Country-income group	Income classification according to World Bank (based on GNI/capita thresholds) Ordinal (low, lower-middle, upper-middle, and high-income) From World Bank World Development Indicators	194 countries
State capacity	Composite indicator of the perceptions of the quality of public services, civil service, policy formulation, implementation, and the credibility of the government's commitment to such policies. Continuous indicator, measured as a percentile rank indicating the country's rank among all countries covered by the aggregate indicator, with 0 corresponding to lowest rank, and 100 to highest rank. Originally developed by Kaufman et al. (2011), obtained via World Bank API	194 countries
Distance between potential TIA partners	Great circle distance between country economic centers Continuous, in km (1 – 19,951) From Head et al. (2020), accessed via Centre d'Etudes Prospectives et d'Informations	194 countries
Common border between potential TIA partners	Binary indicator of sharing a common border Binary From Head et al. (2020), accessed via Centre d'Etudes Prospectives et d'Informations	194 countries
Whether potential TIA partners share a common language	Binary indicator of sharing a common language Binary From Head et al. (2020), accessed via Centre d'Etudes Prospectives et d'Informations	194 countries
Implementation of NCD policies not targeting unhealthy commodities	Number of NCD policies not targeting unhealthy policies that achieved partial/ full implementation Continuous From Allen et al. (2020, 2021)	194 countries

Notes: a: all variables are lagged by one year, with the exception of international political integration, for which we lag the variable by 2 years in 2020 as no data are available in

2019.