Balancing Sovereignty and Cooperation:

The Logic of Deferred Contracting in International Agreements*

Leopoldo Biffi†

November 2025

Abstract

International institutions are known to elicit cooperation based on their stringency, clarity, and enforcement. Yet, it is puzzling that states regularly enter agreements whose implications they know little about. Twenty-first century trade faces major obstacles in regulatory barriers, like national product standards, which raise firms' compliance costs and slow-down trade flows, while institutional commitments to dismantling them remain vague both at the WTO and in preferential trade agreements (PTAs). I introduce the notion of deferred contracting (DC) to theorize how states and pro-trade firms can cooperate on costly international commitments after treaty ratification, so as to bypass domestic vetoes. Examining the case of PTAs, I argue that resorting to DC during implementation leads PTA members to converge their domestic standards governing traded products, including in policy areas that are politically sensitive domestically. I introduce a novel dynamic measure of distance in product regulation built using principal component analysis (PCA), and covering 90 countries between 1995 and 2020. Difference-in-differences estimation reveals that bilateral PTAs with more institutionalized DC fora reduce regulatory distance between members states more than other agreements. I complement my causal estimation with a case study on the EU-Japan Economic Partnership Agreement (EPA). Existing literature has long argued that cooperation revolves around visible commitments formalized during politicized negotiations. This study reveals that, in fact, the politics of implementation is no less fruitful in producing new policy decisions in the shadow of incomplete contracts.

*I wish to thank the following individuals for the precious inputs that have helped improve this work: Niccolò Bonifai, Ricky Clark, Christina Davis, Cédric Dupont, Jeffry Frieden, Robert Gulotty, Connor Jerzak, Doeun Kim, Minju Kim, Dillon Laaker, Quan Li, Siyao Li, Naija Liu, Hannah Loeffler, Ruofan Ma, Jean-Frédéric Morin, Layna Mosley, Harry Oppenheimer, Erica Owen, Evgeny Postnikov, Camille Reverdy, Sung Min Rho, Hongyi She, Stefanie Walter, Hye Young You, Cartland Zhou.

†Princeton University.

1 Introduction

Why do states enter agreements whose implications they know little about? The literature has looked at institutions as means whose ends are clear to their creators (Abbott et al., 2000; Koremenos, 2005; Koremenos et al., 2001; Rosendorff & Milner, 2001). Yet, there seem to be numerous instances of states entering extremely consequential agreements while lacking clarity about the commitments they sign onto. For example, following the United Kingdom's (UK) vote to leave the European Union (EU), the UK and the EU reached an agreement setting the terms of their future relationship, but excluded from it core political decisions like the issue of custom checks between Northern Ireland and the Republic of Ireland. Since the Good Friday Agreement (1998), the seamless flow of goods across the border has been key to maintaining peace in the region. Still, the UK and the EU fully agreed on how this would be maintained only years after the Brexit agreement was adopted, through an informal ex post arrangement known as the Windsor Framework (2023).

I look at the case of regulatory convergence between international trade partners as a hard test for international cooperation in the twenty-first century. Regulatory gaps posed by differing product standards represent a major stumbling bloc for trade liberalization. Their governance complexity and salience has weakened institutional commitments to harmonize, both at the level of the World Trade Organization (WTO) and preferential trade agreements (PTAs). First, product regulations are complex domestic rules whose opacity makes them hard for negotiators to address in the scope of a single agreement (Kono, 2006). Second, regulatory barriers to trade have largely replaced tariffs as new protectionist measures appealing to powerful business groups (Gulotty, 2020; Perlman, 2023). Third, product regulations concerning public health, product safety, the environment, and consumer protection are salient instruments of public-good delivery, which makes them politically costly to retrench through international agreements.

I argue that even incomplete agreements can structure effective cooperation around complex issues like regulatory convergence by establishing institutional spaces for what I call *deferred contracting* (DC). DC consists of states bargaining over new commitments or clarifying existing ones once an agreement is concluded. DC occurs during implementation, beyond the boundaries of formal treaty clauses but still in the shadow of a signed agreement which provides the principal institutional structure for members' interactions. DC takes place through treaty-based joint committees or other fora, where state bureaucrats and interest groups coordinate policies that are relevant to the implementation of an agreement, define what counts as compliance, and can modify certain commitments ex-post, bypassing the close domestic oversight typical of open negotiations.

This work shares some good news about international cooperation and a case for the power of institutions

in global governance. It contends that institutions first lacking clear or binding commitments can still provide their members with structured points of contact through which more meaningful cooperation can follow. Looking at all bilateral preferential trade agreements (PTAs) negotiated since 1995, I explore whether the presence of institutional spaces for DC within PTAs enhances convergence in domestic regulatory standards between treaty members. I introduce a novel measure of regulatory distance, computed through principal-component analysis (PCA), for 90 countries between 1995 and 2020. Relying on difference-in-differences estimation with propensity-score weighting (Imai et al., 2023), I find that PTAs setting up a greater number of institutions for DC, like treaty joint bodies and committees, reduce regulatory distance between member countries more relative to signatories of other agreements, as well as in absolute terms. Results are robust across various model specifications and country dyads. Relying on a mixed-method approach, I complement the above results with a case study on DC dynamics that followed the ratification of the EU-Japan Economic Partnership Agreement (EPA), in 2019, with a focus on cooperation around sanitary and phytosanitary measures (SPS).

While the literature has long established that institutional design varies greatly in strength, precision, and enforceability (Abbott et al., 2000), we still know little about the tangible effects that lowly-legalized commitments have on interstate cooperation. We know that, by design, incomplete contracts grant policy-makers greater decisional flexibility in the long term (Abbott & Snidal, 2000; Goldstein & Martin, 2000). These considerations tell us little, however, about whether looser commitments have an effect on international cooperation—and if this is a positive one. A theory of deferred contracting aims to fill this gap by investigating and testing how regulatory convergence can happen in the shadow of however vague contracts. In so doing, I also revisit the nexus between domestic politics and international negotiations. While scholars have emphasized how the content of international agreements is constrained by domestic politics (Putnam, 1988; Tsebelis, 1995), this paper shows how implementation can advance international cooperation beyond what would be negotiable during open bargaining, when domestic veto playing is at its maximum. This work also advances our understanding of incomplete contracting in the context of international agreements (Hart & Moore, 1988; Maggi, 2011). Incomplete-contracting literature has overlooked how contract incompleteness can actually further cooperation, by allowing costly commitments to be worked out ex-post. This seems to be a growing phenomenon as trade cooperation becomes more technical and politicized, both aspects that transfer considerable policy discretion to actors involved in implementation (Bögner, 2024; Claussen, 2024; Czechowska, 2022; Dür & Gastinger, 2023). More broadly, this study offers a first empirical examination of whether PTAs can explain variation in levels of regulatory harmonization between states. This phenomenon has been discussed widely in the literature but its empirical testing has lagged due to challenges in modeling regulatory distance efficiently (Maggi & Ossa, 2021). In sum, different levels of cooperation are not simply

a function of changes in formal obligations. They are shaped by the nature and frequency of long-term interactions, the (dis-)activation of domestic veto points, and the ability of institutional venues to adapt to evolving contingencies.

2 Analytical Framework: Trade Politics and Deferred Contracting

What allows governments to achieve free trade when doing so clashes with the interests of their constituents? I propose that trade negotiators will aim for the strongest liberalizing agreement that does not encounter the veto of constituents. For governments, making manifest their intention to liberalize can be politically costly, exposing an agreement to the risk of failed ratification (Biffi, 2020). As Goldstein and Martin (2000) observe, '...legalization that involves highly precise and transparent rules can have the unintended effect of encouraging the mobilization of protectionist forces that see themselves as probable losers from an agreement' (606). This is not just true for agreements lowering tariffs but increasingly so for trade commitments to harmonize with international product standards, which lowers barriers to the entry of foreign products. To avoid ratification costs, policymakers can set the range of policies that are decided during open negotiations and those deferred to implementation. During open bargaining, costly agenda items attract the attention of harmed constituents who can leverage their veto powers to demand modifications of the agreement or otherwise impede ratification (Putnam, 1988; Tsebelis, 2011). At implementation, in contrast, successful ratification has already de-activated most veto powers, making liberalizing decisions less scrutinized.

Deferred contracting (DC) allows two or more governments to take advantage of lower public scrutiny during the implementation of an international agreement. Similarly to delegation to international organizations (Vaubel, 2006), DC removes policy decisions from the direct control of voters and parliaments as policies are developed behind closed doors. In so doing, DC is also distinct from passive enforcement as it entails *new* decisions being made after treaty adoption. For example, during negotiations of the Comprehensive Economic and Trade Agreement (CETA), European Union (EU) and Canadian officials decided to leave off the negotiation agenda the issue of maximum residue levels (MRL) of pesticides in agricultural products. The EU faced significant pressure from its trade partners, including Canada, to loosen its MRL regulation for several chemicals, which could have limited the entry of foreign products. At the same time, European NGOs like the Pesticide Action Network had the issue under the radar, lamenting how too permissive MRL standards could harm public health in the EU (PAN Europe, 2015). MRL decisions were kept out of CETA's text and were tabled during "quieter" talks that resumed only during the implementation of the agreement (European Commission, 2018b).

The room for DC increases when international agreements are more incomplete. Incomplete-contracting

problems are prevalent in modern trade agreements, where complex rules on competition, intellectual property, product standards, and environmental protection have expanded the agenda beyond what used to be primarily about tariff cooperation (Baccini, 2019; Maggi, 2011; Maggi & Ossa, 2021). Domestic standards regulating traded products have emerged as a major obstacle to globalization insofar as these rules differ cross border, placing fixed compliance costs for exporters that are complex to quantify and govern (Kono, 2006). Implementing trade deals takes longer than it used to when commitments primarily revolved around tariffs and treaties no longer work as precise legal scripts to follow: treaty commitments need to be interpreted, domesticated, and translated into operative rules which can expand, modify, or retrench the original international commitment. For example, the US implementing legislation for the United States–Mexico–Canada Agreement (USMCA) does not extensively codify commitments made in the agreement, limiting the oversight powers that Congress can exercise over implementation. This lack of precision increases the discretion of the executive bodies over the operationalization of the agreement. Federal departments and agencies can selectively focus on implementing certain aspects of the agreements while reneging on other commitments made (Claussen, 2024). DC allows treaty members to continue making decisions that clarify, expand, and modify existing commitments under an agreement that is already adopted, decisions which, if tabled during open negotiations, would likely jeopardize the ratification of an agreement.

The criticality of DC is revealed when measuring the lobbying traffic occurring after the ratification of international agreements. We know that a substantial amount of interest-group influence is exercised only after new legislation is passed (You, 2017). I emphasize how this logic can be extended to incomplete international agreements like PTAs. Contract incompleteness keeps incentives to lobby high as several treaty commitments are yet to be defined and operationalized. Figure 1 (below) displays the number of lobbying reports filed per yearly quarter to the US Congress and concerning the Implementation Act of the USMCA, which was ratified between December 2019 and January 2020. The plot reveals that, although the quarterly lobbying traffic decreased post-Q4 of 2019, 70% of all measured lobbying activity around the USMCA still took place only after the agreement's ratification.

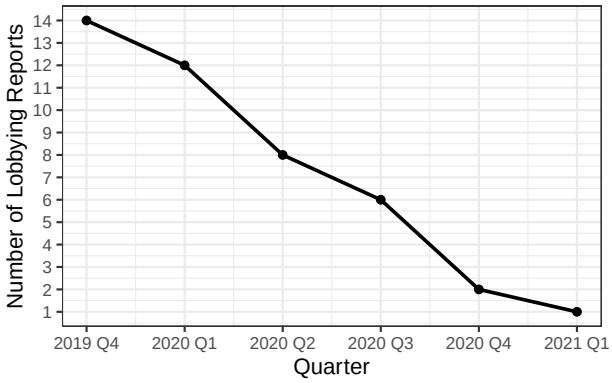


Figure 1: Intensity of lobbying activity around the USMCA Implementation Act. The ratification cutoff is Q4 of 2019. No lobbying entry was detected prior to Q4 of 2019 and after Q1 of 2021. The data are retrieved from LobbyView (Kim, 2018).

3 The Choice to Defer Contracting

Governments resort to DC as a negotiation strategy to leave parts of international agreements opaque, and and to work them out post ratification. This can be due to either the complexity, time intensiveness, or the political salience of such commitments if finalized or publicly exposed during open bargaining. I argue that the choice to defer contracting is a function of the gap between the cooperation goals set by international negotiators and the level of cooperation acceptable by domestic constituents. No space for DC exist if the preferences of international negotiators and domestic constituents perfectly overlap, which would allow for a precise agreement to lay out each and every commitment, succeed ratification, and be passively implemented. The value of DC grows as the size of the international-domestic win-set shrinks, that is as the number of international cooperation commitments within an agreement that are politically untenable domestically grows (Putnam, 1988). This is true both if international negotiators favor more (less) cooperation and constituents less (more) cooperation. In the former case, DC helps international negotiators advance cooperation beyond the baseline of a treaty text, ex post. In the latter case, negotiators use DC to cater to special interests not to cooperate when original commitments in the original agreement would set higher levels of cooperation.

For example, the US-Japan PTA (USJTA) was negotiated in 2019 as a a limited-scope agreement focusing primarily on agricultural tariffs and digital trade. Its limited scope punctually catered to the market access demands of US agricultural constituents in the coming up to the 2020 Presidential elections. This "mini deal" naturally left out key chapters like automotive, non-tariff barriers to trade, and trade in non-agricultural goods. Outstanding agenda items were almost entirely delegated to the US-Japan Partnership On Trade,

a forum set up in 2021 under the auspices of the USJTA to govern bilateral trade cooperation. In 2023, through the Partnership on Trade, the U.S. and Japan swiftly concluded a Critical Minerals Agreement (CMA) that was instrumental to qualifying Japan as a U.S. trade partner eligible for EV battery tax credits under the Inflation Reduction Act (IRA). The Partnership on Trade served as a forum for negotiation of the CMA, which was kept as an executive agreement not requiring congressional approval, which would have likely been an obstacle considering Republican opposition to the IRA (Department of the Treasury, 2025).

If deferred contracting (DC) is instrumental to making decisions that go beyond what is written in the original contract, its consequentiality must vary with agreement-design aspects. Otherwise, DC would simply be employed as routine implementation strategy that states resort to because they want any agreement to be successfully implemented. I put emphasis on the length of an agreement's agenda and its depth as important scope conditions shaping incentives for the parties to choose to defer contracting over some items, by setting up ad-hoc institutional bodies. First, a longer agenda is more likely to carry specific issue items that can encounter the opposition of different constituents if fully exposed during open bargaining. While a longer agenda has been found to ease negotiation breakthroughs, it does so precisely by not isolating potential focal points of contention and thus broadening the constituency base in support of the agreement (Davis, 2004). As a flipside, this also leaves a number of politically salient items in need to be worked out ex post. Second, deeper agreements are also higher in salience, for they more strongly 'require states to depart from what they would have done in [their] absence' (Downs et al., 1996, p. 383). Several economic agreements, like PTAs, have deepened over the past two decades, coming to cover a growing number of sensitive domestic policies including product standardization, intellectual-property rules, market competition, environmental law, whose coordination entails lengthy decision-making processes (Baccini, 2019; Maggi & Ossa, 2021). In sum, a tradeoff between treaty precision and political discretion shapes incentives for states to resort to DC.

Looking at variation in treaty length is a useful heuristic showing that DC is not a default negotiation feature, but is shaped by governments' political incentives to continue negotiating. To illustrate this intuition, Figure 2 (below) considers all international environmental agreements (IEAs) concluded between 1945 and 2015, and shows the frequency of IEAs based on their textual length and whether or not they establish an intergovernmental committee (e.g., a conference of the parties) serving as a forum for dialogue, decision-making, and information-sharing after adoption.

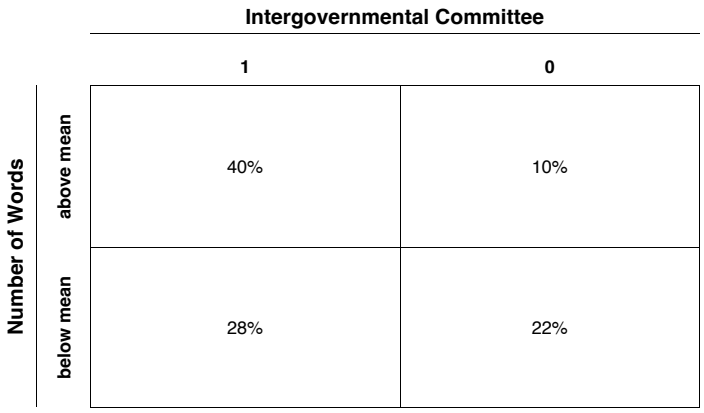


Figure 2: Percentage of IEAs establishing an intergovernmental committee based on the IEA length (number of words). The data ar retrieved from Laurens et al. (2023) and Morin et al. (2024).

While about 70% of IEAs establish a committee, the distribution varies markedly by treaty length. For IEAs below average length, the shares are 28% with a committee versus 22% without (of total agreements). For IEAs above the mean, the gap widens significantly, with 40% of treaties with a committee versus only 10% without one (of total agreements).

In trade, DC has become critical as commitments in PTAs have expanded towards governing regulatory convergence, becoming obfuscated by lengthy processes of domestic institutional adjustment (Kono, 2006). In the field of technical barriers to trade (TBT) and sanitary and phytosanitary standards (SPS), which constitute the lion's share of trade regulatory barriers, several PTAs have produced a vast architecture of fora structuring DC. TBTs are national measures requiring products to comply with given safety, health, environmental or security standards. SPS are analogous to TBTs, but are specific to food safety and pest control. PTA institutions for TBT and SPS governance include both administrative and consultative bodies, joint committees able to issue policy decisions and binding recommendations. Figure 3 (below) plots the average count of DC institutions for TBT and SPS across per PTA over time.

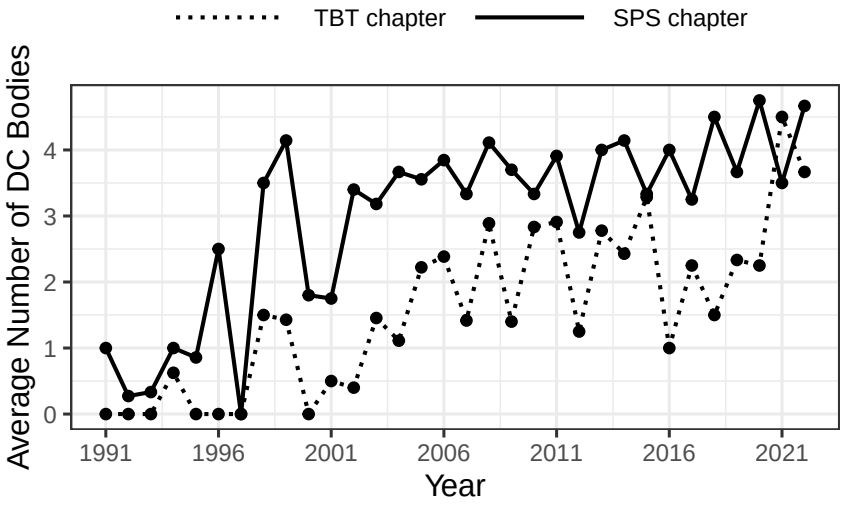


Figure 3: Average yearly count of DC bodies governing the TBT and SPS chapters in PTAs negotiated between 1990 and 1991. Data are retrieved from the World Bank Dataset on Deep Trade Agreements (Rocha et al., 2021).

TBT and SPS chapters have evolved from including on average few to no institutional bodies to agreements producing the full spectrum of DC institutions: joint committees carrying official decisional and amendment powers as well as consultative and administrative bodies that meeting regularly. The growing articulation of these bodies suggests an increased specialization of their governing tasks. In EU PTAs, for instance, specialized committees like those on SPS and TBT, report to a single higher-level committee in charge of ministerial-level decisions adopted post-ratification. This degree of articulation allows these bodies to operate as self-sufficient governance apparatuses able to systematically administer the life of an agreement. This is indicative of how DC plays a more regularized governance function than single dispute settlement bodies, which only come into play in the event of a formal dispute between members, or when tasked to rule over a party's exercise of its treaty rights, like invoking flexibility clauses (Rosendorff, 2005). Furthermore, PTA joint committees are not an instance of policy delegation. Indeed, they often gather the same officials and bureaucrats involved in negotiations (Bögner, 2024; Dür & Gastinger, 2023). In sum, the key aspect setting DC apart from regular dispute-settlement bodies is that the former make *new* policy decisions, rather than being delegated to interpret existing ones.

4 Deferred contracting and Regulatory Convergence

Having established how DC can be institutionalized, I elaborate on how it can entail greater regulatory convergence following the ratification of a PTA. I argue that DC enhances regulatory convergence by intensifying policy coordination between counterpart governments on politicized issues that would expose an agreement to ratification obstacles if tabled before ratification.

Regulatory convergence presents a hard test for international trade cooperation. Its politics is underpinned by a different set of considerations from traditional tariff-based concerns pitting import-competing vs. export-oriented groups. First, regulatory convergence is encouraged mainly by large multi-national corporations (MNCs) lobbying for standards harmonization across specific regions where they operate (Young, 2016). This can enhance the control MNCs enjoy over specific markets, erecting compliance barriers to the entry of smaller and less productive firms from outside (Gulotty, 2020). Second, the ambivalent nature of regulatory standards as both public-good enhancing and trade restrictive leads two main types of domestic groups to oppose regulatory convergence: citizens fearing that regional convergence would bring national regulatory "chills" and lobbying groups that align with these concerns for purely protectionist reasons (Perlman, 2023). For example, the agricultural lobby Japan Agricultural Cooperatives (JA) sustains Japan's articulated SPS measures as market barriers, rather than under a mandate to uphold public health (Suzuki, 2017). Third, institutions like the WTO TBT and SPS Agreements, as well as PTAs, lack the precision of tariff-line liberalization, limiting themselves to softly mandating alignment with internationally recognized standards. Similarly loose commitments make it difficult for treaties to formally prescribe standards parties shall converge to and for which products or services.

Figure 4 (below) shows trends in the average levels of regulatory distance (red line) between country dyads part of a bilateral PTA over the last 3 decades. This regulatory distance measure captures the absolute distance, in each year, between two countries' regulatory ideal points, which are computed dynamically running principal component analysis (PCA) over frequencies of national regulatory measures by product type. [1] The plot suggests that even as domestic product standards have generally diverged (the average regulatory distance increased by about 1%), states have continued to enter PTAs.

[1] Section 5 discusses the construction of this measure.

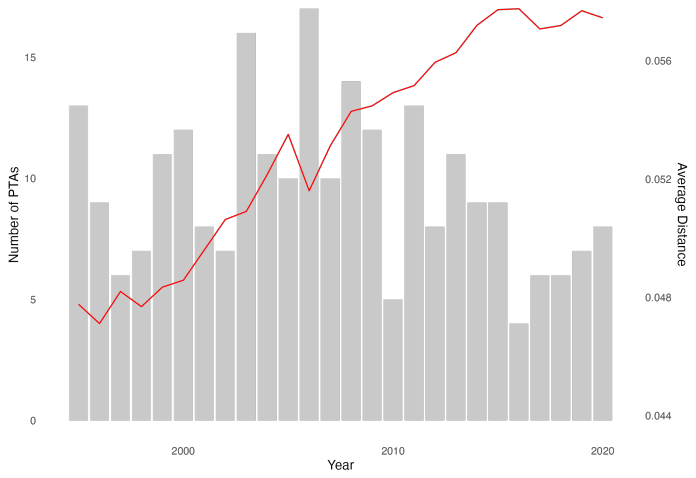


Figure 4: Regulatory distance and number of PTAs over time.

PTAs, especially modern ones, are largely understood as having little to do with tariff liberalization (a task largely up-taken by the WTO) but rather regulatory cooperation (Baccini, 2019; Maggi & Ossa, 2021). Yet, it seems that not all PTAs do equally well at enhancing regulatory convergence, despite them loosely mandating adherence to the same internationally recognized product standards and echoing the same WTO Agreements governing regulatory cooperation.

As discussed, the ratification of a treaty fundamentally changes the bargaining position of member governments. While the veto powers of constituents, leveraged at ratification, require negotiators to simultaneously contract with their international counterparts and domestic constituencies over a mutually satisfying policy outcome (Putnam, 1988), ratification largely shields subsequent policies from domestic constraints. I hypothesize that specific liberalization commitments which, in an open negotiation setting, may encounter the opposition of various stakeholders, can be attained by resorting to DC. The institutional gaps left by incomplete contracting grant trade officials greater policy autonomy from their domestic principals. I put emphasis on the design of institutions for DC as decision-making fora, rather than mere enforcement or flexibility tools (Hart & Moore, 1988).

I shall now discuss how the institutionalization of DC can aggregate political interests towards regulatory convergence. In particular, why would DC necessarily lead a PTA to enact greater market integration? I contend that this is the case as government-level mutual requests for greater market access will not be credibly challenged by protectionist pressures, in the aftermath of ratification and in the absence of credible veto powers. At the international level, requests from government officials within PTA joint committees voice the demands of their exporting businesses for greater market access. On regulatory policies like TBT and SPS, this entails either harmonizing required standards with those of the exporter or recognizing the

exporter's standards and product verification procedures as equivalent to domestic ones.

While citizen groups are generally understood as having limited lobbying access, they can successfully resort to outside-lobbying tactics by politicizing specific negotiation items that gain the attention of veto players like parliaments (Dür & Mateo, 2016). For example, during the negotiation of CETA, citizens groups focused heavily on the issue of investor-state dispute settlement (ISDS) contained in the original agreement, pushing the European Parliament to demand the removal of ISDS provisions from the final agreement before ratification (Biffi, 2020). In contrast, during implementation, pro-liberalization firms willing to harmonize domestic standards are the most represented and active group. A key asset for these pro-trade coalitions during implementation is access. Businesses tend to channel their lobbying efforts through direct and repeated contacts with trade officials in charge of implementation (Dür & Mateo, 2016).

Besides these structural advantages, sunk-cost considerations will drive pro-liberalization coalitions to push their governments towards greater convergence. Businesses that self-select into PTA lobbying are mainly large multinational companies that see these institutions as instrumental to regional regulatory harmonization, enabling them to streamline compliance requirements and enhance their market and value-chain presence (Bombardini, 2008; Maggi & Ossa, 2020). As successful lobbying for ratification entails considerable upfront costs, businesses having already committed resources to mobilizing in favor of an agreement will be keen to reap the greatest market returns on their lobbying investment. They will demand that PTA implementation structures concrete steps towards regulatory convergence, especially in the presence of loose and uncertain contract-level obligations. Indeed, the literature has found that higher economic-policy uncertainty prompts already active firms to scale up their lobbying efforts (Shang et al., 2021). In this light, I expect business lobbyists to influence post-ratification decisions in ways that either reinforce regulatory convergence commitments already enshrined in the agreement or extend convergence above the PTA baseline.

In contrast, I expect the lobbying power of anti-convergence coalitions to considerably diminish at implementation. At one level, this is related to the lack of ratification powers already discussed. For example, several European NGOs which had vocally lamented the environmental consequences of the CETA have become less active since the agreement's provisional ratification in 2017, redirecting their lobbying efforts to open negotiations like those between the EU and the MERCOSUR. This is coherent with findings that citizen groups like environmental NGOs tend to pressure lawmakers, whose power greatly diminishes after ratification. At another level, the interest of these actors in actively allocating their (limited) resources to ex-post lobbying will also diminish. As discussed, the adoption of a PTA itself already entails higher levels of liberalization and regulatory convergence than average. In this context, the marginal effect that counter lobbying is expected to have in reverting the convergence effects of a PTA is small, and not rationally deserving of additional lobbying resources. At best, the tactics of protectionist groups, especially import-

competing businesses opposing regulatory convergence, would be aimed at delaying, rather than reverting these processes of regulatory convergence. For example, as the case study elaborates on, Japan Agricultural Cooperatives (JA) has sought to delay more streamlined recognition of EU agricultural products by Japan ever since the enforcement of the EPA. Overall, I expect PTAs producing articulated institutions for DC to lead to greater regulatory convergence between their members than agreements that tend to concentrate decision-making before ratification. I formulate the following hypothesis:

Hypothesis: *PTAs establishing more institutions for deferred contracting will lead to greater regulatory convergence between their member states.*

5 Measuring Regulatory Distance and Convergence

I test the effect of signing PTAs on a novel dynamic measure of *regulatory distance* between pairs of signatory countries, from 1995 to 2020 included. Yearly decreases (or increases) in this distance measure allow to gain a sense of the amount of convergence (or divergence) experienced by country dyads over time. The measure covers the following regulations: technical barriers to trade (TBTs), sanitary and phytosanitary standards (SPS), and pre-shipping inspections (PSI). PSI consist of inspections of goods and other formalities that exporting countries are required to perform in order for the goods in questions to qualify for access to certain foreign markets. TBT and SPS pose by far the largest amount of non-tariff barriers in modern trade in goods.

5.1 Data on NTMs

I collect data on regulatory measures from the UNCTAD Trade Analysis and Information System (TRAINS) Database (UNCTAD, 2024). TRAINS records TBT, SPS, and PSI measures, by country, by year. I consider 90 countries in the 1995-2020 period. The EU is taken as a single country as the establishment of the European Single Market (1986) preceded the negotiation of GATT/WTO TBT and SPS rules at the Uruguay Round. Each measure is labelled according to the UNCTAD Multi-Agency Support Team (MAST) taxonomy of non-tariff measures, which is the most systematic international classification criterion for NTMs (UNCTAD, 2019). My data displays the number of NTMs cumulatively in place per country in any given year by product type. Product types are considered at the two-digit level of the harmonized schedule (HS), which consists of 96 chapters. Each chapter represents a family of products, like vegetables, cereals, or ma-

chinery. The TRAINS dataset is broken down into individual yearly matrices where each row corresponds to a country, and columns to the 96 HS product families. Values in the matrix grid are the cumulative sum of NTMs in place under each product family in a given country in a given year. I run principal component analysis (PCA) on this data matrix to extract a composite country-year measure of regulatory density by product type. Table 1 (below) displays the 22 measure types included in my data.

Measure Type	**UNCTAD MAST Code**	**NTM Type**
Authorization requirements for importing certain products	B14	TBT
Authorization requirements for importers	B15	TBT
Labeling requirements	B31	TBT
Marking requirements	B32	TBT
Testing requirements	B82	TBT
Certification requirements	B83	TBT
Inspection requirements	B84	TBT
Traceability requirements	B85	TBT
Prohibitions for sanitary and phytosanitary reasons	A11	SPS
Geographical restrictions on eligibility	A12	SPS
Systems approach	A13	SPS
Authorization requirement for sanitary and phytosanitary reasons for importing certain products	A14	SPS
Authorization requirement for importers for sanitary and phytosanitary reasons	A15	SPS
Labelling requirements	A31	SPS
Marking requirements	A32	SPS
Packaging requirements	A33	SPS
Testing requirements	A82	SPS
Certification requirements	A83	SPS
Inspection requirements	A84	SPS
Direct consignment requirements	C2	PSI
Requirement to pass through specified port of customs	C3	PSI
Import monitoring, surveillance and automatic licensing measures	C4	PSI

Table 1: Measure types included in the data (UNCTAD MAST classification).

5.2 PCA: Dynamically Measuring Countries' Regulatory Distance

I use principal component analysis (PCA) to create composite country-year variables capturing core characteristics in national patterns of adoption of TBT, SPS, and PSI measures. I call this a country's *regulatory ideal point*. The intuition is simple. By examining the number of TBTs, SPS, and PSI regulations a country has in place across the 96 product families at different points in time, one can dynamically measure which goods and related risk factors a country prioritizes regulating at different points in time. PCA is used to project high-dimensional data onto a low-dimensional space (the principal subspace). It operates by reducing the number of variables in an original data matrix by building composite variables from existing ones and ensuring that each of the new ones are independent from each other, i.e. that each of them summarizes

new information from the original data. The first principal component (PC1) is the orthogonal projection of data-points which reflects highest share of variance among units, in our case countries. PCA is a widely employed computational technique in the social sciences, which has been used to build various composite variables by shrinking more complex original dimensions. These include measures of political culture, institutional instability, presidential power, party-system characteristics (Berggren et al., 2012; Doyle & Elgie, 2016; Jackman & Miller, 1996; Magyar, 2022).

The shrinkage of the original data operated by PCA can also be illustrated by visualizing the rotation of the data matrix. PCA reduces the original dimensions in a dataframe by mathematically rotating the original data-matrix in a way that maximizes the variance between units. Each one of the new rotated dimensions represents a new composite variable of the original dimesions and each is independent from one another. Figure 5 (below) shows how a dataset containing 96 dimensions of interests, in our case product families for each of the HS chapters, can be reduced to a single dimension of interest (PC1) capturing most information about a country's regulatory intensity.

	Live animals (HS-1)	Vegetables (HS-2)	...	Miscellaneous (HS-96)
Country A	4	20	...	20
Country B	1	3	...	12
Country C	8	2	...	7

→

	PC1
Country A	0.20
Country B	1.5
Country C	2.6

Figure 5: Illustration of the data shrinkage operated by PCA. In each given year, the data-matrix containing a country's frequency of regulatory measures in place for each of the 96 product families (columns) is reduced to a single PC1 score. Numbers in the tables are purely illustrative.

The distance computed between the PC1 scores of countries, taken pairwise, is the most informative for the purpose of this study. In each year, this distance shows how far apart countries are along the PC1 line, considering them pairwise in all possible combinations. For the 1995-2020 period, the final dataset contains the distance between the PC1 scores all 90 countries in the dataset combined pairwise. Distance values are scaled from 0 to 1 to ease the interpretation of the causal estimation that will follow.

Let

$$d_{ij,t} = \text{PC1 distance between countries } i \text{ and } j \text{ in year } t \tag{1}$$

where $t \in [1995, 2020]$, $i \neq j$, and all countries are considered pairwise. The scaled distance is given by

$$\tilde{d}_{ij,t} = \frac{d_{ij,t} - \min(d)}{\max(d) - \min(d)} \in [0, 1], \tag{2}$$

PCA also learns about additional directions explaining shares of data variance in decreasing order from PC1. The second principal component (PC2) represents the line onto which datapoints are projected with the second largest variance, and so on: $PC3...PC_n$. An important feature of these additional components is that these are orthogonal to the PC1 and to one another: $\text{PC}_1 \perp \text{PC}_2 \perp \text{PC}_3 \perp \ldots$. This ensures that each additional dimension explains a share of data variance that is not already captured by previous components. In the causal estimation that follows, regulatory distance is computed exclusively between PC1s as this dimension alone captures over 50% of total variance in the data. [2]

5.3 Regulatory Distance: Descriptive Evidence

To get a sense of how reliably PC1 measures regulatory distance, one can examine how this measure changes across countries over time. At a macro level, global regulatory distance seems to have grown by around 1% since the establishment of the WTO in 1995. That is coherent with general findings in the literature that the lion's share of regulatory activity in trade is found in this period (see Figure 4 above). Despite the Uruguay Round having strengthened the regime for regulatory cooperation in trade, mainly through the WTO TBT and SPS Agreements, the broad interpretability of obligations not to restrict trade via national standards left states ample room to regulate products and services (Gulotty, 2020). Additionally, the weak enforceability of TBT and SPS multilateral rules failed to deter countries' unilateral pursuit of regulatory barriers (Barton et al., 2008). For example, in the late 1990s, the EU adopted a general import ban on genetically-modified products (GMOs), further strengthening its regulatory barriers on agricultural imports. In the 1980s the EC had already enforced a ban on imports of growth-hormone treated beef, which is still in place to this day despite having been deemed illegal by a WTO Appellate Body ruling in 1998.[3]

Figure 6 (see p. 17) plots PC1 scores for the 90 countries in the data over the past 3 decades. Over this period, new regional markets have been formed and existing ones have been strengthened, expanding institutionalized regulatory cooperation, mainly through far-reaching PTAs or more ad-hoc mutual-recognition agreements (MRAs) for product and sanitary standards. The four world plots generally indicate that the majority of countries increased their regulatory activity between the mid-1990s and the late 2000s. By observing over time the evolution of regional trade blocs, one can best understand how reliably PC1 captures the proximity

[2] See Appendix, Section B.

[3] Appellate Body Report, European Communities: Measures Concerning Meat and Meat Products (Hormones), WT/DS26/AB/R, adopted 13 February 1998.

among closely integrated trade partners and the greater distance between lesser integrated ones. Intuitively, more integrated markets should have a more similar cross-product regulatory activity, and color gradients, and vice-versa.

Looking at the European continent, the plots indicate that close external partners of the EU have pursued regulatory alignment with the Union over the years. For example, Switzerland, whose PC1 score was visibly lower than the EU's in the mid-1990s, closely aligned with the EU by the early 2000s. This reflects how Switzerland and the EU adopted a series of bilateral agreements in the late 1990s strengthening regulatory cooperation and including, inter alia, MRAs across a large number of industrial goods, as well as a Veterinary Agreement expanding SPS equivalence and eliminating veterinary border checks (Lavenex, 2009). The fact that Switzerland's PC1 score is higher than the EU's in the year 2010 is likely explained by the intense regulatory activity of Switzerland to catch up with the EU framework, including by amending its Technical Barriers to Trade Act to streamline the conformity of incoming EU products.

Considering another European trade bloc, the European Free Trade Area (EFTA)[4], one can see how the PC1 scores of Switzerland, Liechtenstein, and Norway move closely over the years. The maximum regulatory distance recorded between EFTA members amounts to 6% for Norway and Switzerland in 1995, although the two countries have progressively converged, presenting a 1.5% distance by the early 2020s.

Moving to the American continent, the evolution of PC1 scores reflects market integration within the MERCOSUR.[5] In 1995, when MERCOSUR established itself as a customs union, the PC1 scores of the bloc's two largest economies, Argentina and Brazil, were closely aligned, while Paraguay and Uruguay are aligned at a lower color gradient. The 4 countries appear fully harmonized by the year 2000, and remain consistently so throughout the 25 years covered by this study. Bolivia and Chile also show regulatory convergence towards the MERCOSUR from the year 2000 onwards, as they both joined the bloc as associate members in 1996.

Considering Asia, the PC1 measure also appears to reliably capture regulatory alignment within the Association of Southeast Asian Nations (ASEAN). From 1995 onwards, Cambodia, Thailand, Myanmar, and Vietnam are visibly aligned int their PC1 trends, with their PC1 scores co-varying both upwards (1995-2000; 2005-2010) and downwards (2000-2005; 2015-2020). These patterns are coherent with institutional steps taken by ASEAN to strengthen its regulatory cooperation over the years. In 2005, for example, the bloc adopted a far-reaching MRA covering standards for products ranging from telecom and electric equipment to cosmetics and pharmaceuticals. Indonesia, being by far the largest market in the bloc, follows a somewhat more autonomous pattern. While its score closely approaches that of other members between 2000 and 2010, it has evolved differently ever since.

[4]EFTA members include Switzerland, Liechtenstein, Norway, and Iceland.

[5]MERCOSUR comprises Argentina, Brazil, Paraguay, Uruguay as full-fledged members; Chile, Peru, Colombia, Ecuador, Guyana, Suriname as associate members; and Bolivia in the process of completing its accession.

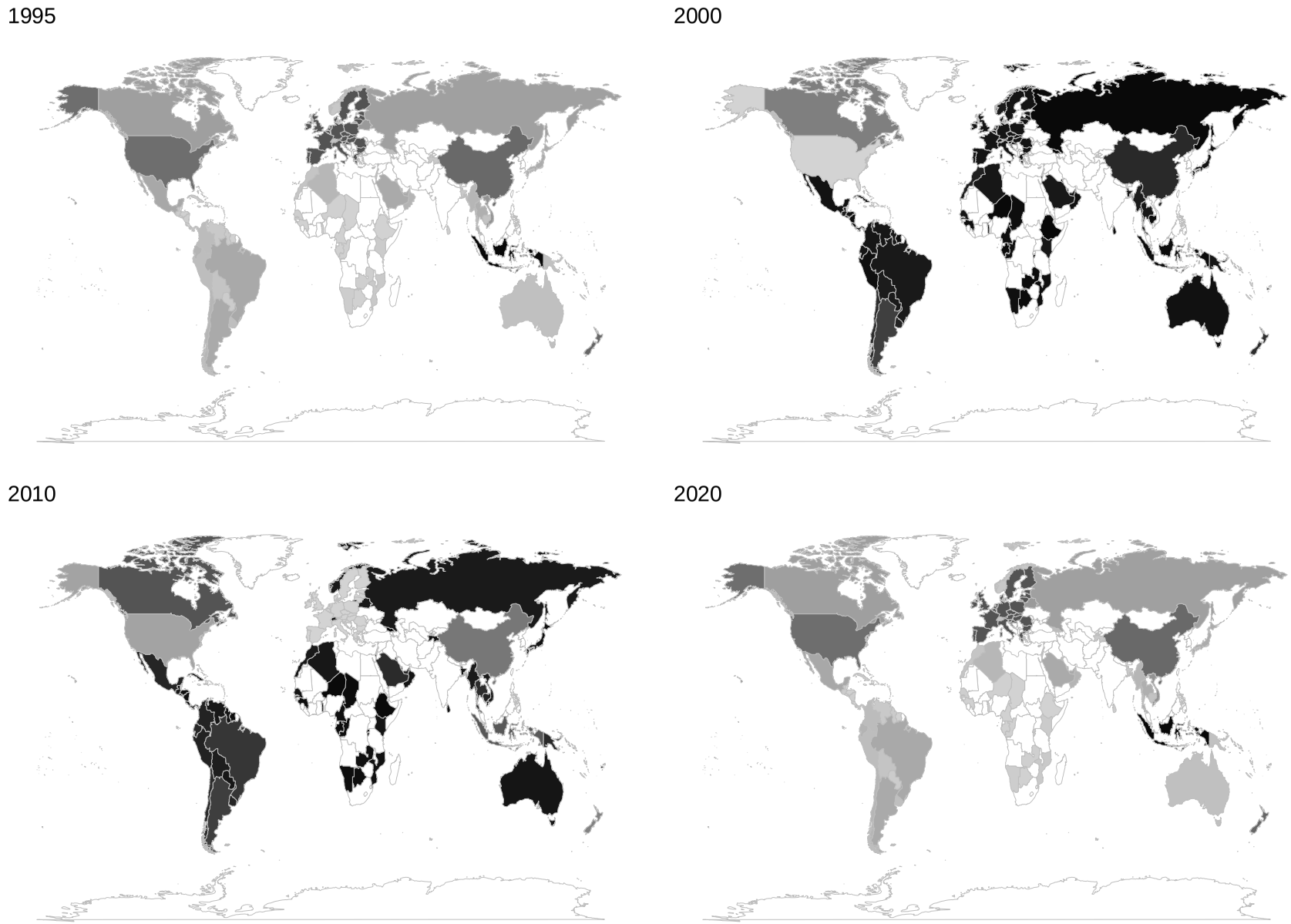


Figure 6: Geographic distribution and evolution over time of PC1 scores. Darker-shaded regions indicate a higher PC1 score, while lighter-shaded ones a lower one, on a scale from 0 to 1. White regions indicate missing data.

Australia and New Zealand also reflect patterns of institutionalized regulatory cooperation over the years. In 1998, the two markets entered the Trans-Tasman Mutual Recognition Arrangement (TTMRA), which introduced a universal adequacy between products from either country. The years following the entry into force of the TTMRA (1998) show a perfect alignment between the PC1s of Australia and New Zealand (2000-2005). This seems to reflect New Zealand's effort to catch up with the adoption of regulations, product registration, and approval procedure already in place in Australia. This process of alignment was institutionally overseen by bodies like the Trans-Tasman Food Standards Council, which later became Food Standards Australia New Zealand (FSANZ), which was tasked with developiong the Australia New Zealand Food Standards Code. From 2010 onwards, the regulatory density of Australia increases to then drop again in 2020. New Zeland's score remains constant and lower in the 2010-2020 decade, coherently with New Zealand being a smaller market than Australia and mainly a receiver of regulation from its larger neighbor.

6 Research Design

I test the effect that the presence of DC institutions in a PTA has on the regulatory distance between its members. As my treatment, I rely on a count variable capturing how many institutions for DC a PTA produces in the field of TBT and SPS. Dyads entering PTAs whose number of DC institutions is above average count of institutions are treated. I rely on a difference-in-differences (DiD) design leveraging the timing of agreement signature. I test the effect of singing a PTA on an unbalanced panel of undirected country dyads (the units of analysis) for the period between 1995 and 2020. I rely on the "PanelMatch" estimator developed by Imai et al. (2023), which allows to weight treated and untreated units on a series of relevant covariates, to ensure that they have nearly identical characteristics in the time lag L before treatment assignment (4 years). This helps ensure that DiD's assumption of parallel trends is satisfied. I refine the treatment and control groups through propensity-score weighting. I then compute the DiD estimator as formalized by the following equation:

$$\hat{\tau}(F,L) = \frac{1}{\sum_{i=1}^{N}\sum_{t=L+1}^{T-F} D_{it}} \sum_{i=1}^{N}\sum_{t=L+1}^{T-F} D_{it} \left\{ (Y_{i,t+F} - Y_{i,t-1}) \sum_{i' \in M_{i,t}} \omega_{i'}^{i} t (Y_{i',t+F} - Y_{i',t-1}) \right\}$$

If a country dyad signs a PTA whose number of DC institutions is above average, D_{it} is equal to one, whereas $\omega_{i'}^{it}$ is the weight assigned to the respective control groups. The final estimate $\hat{\tau}$ expresses the average treatment effect for treated units (ATT). The ATT represents the average of the DiD estimators computed across each ps-weighted set. In this design, ATT estimates are reported for the six-year period

(t+6) following PTA signing. For the testing of my hypothesis, I estimate the effect that entering a PTA with highly institutionalized DC fora has on the regulatory distance between its members. My estimation considers the population of country dyads which entered a PTA between 1995 and 2020, covering a total of 90 country dyads. This allows to capture whether entering PTAs with more articulated DC institutions affects regulatory convergence any differently from entering just any other PTA.

The treatment is assigned in the year of PTA signature, and consists of a dummy variable capturing whether the count of DC institutions established by a PTA in the field of TBT and SPS is above or below average. The average is computed on a count variable (*Dc.Institutionalization*) comprising 18 dimensions of DC institutionalization for TBT and SPS, as coded in the World Bank Dataset on Deep Trade Agreements (Espitia et al., 2020; Stone & Casalini, 2020). It is computed across a total of 283 PTAs. Table 2 (below) displays the counting and weighing of treaty-design dimensions this treatment variable.

Rationale	**Count**
TBT	
Is a regional body established?	+1
Is there a regional dispute settlement body?	–1
Are there regional consultations foreseen to resolve disputes?	+1
Is there a regional dispute settlement body for TBT?	–1
Is there a mechanism to issue recommendations?	+1
Are recommendations mandatory?	+1
Is the recourse to the dispute settlement for technical regulations disallowed?	+1
SPS	
Do the parties establish SPS contact/inquiry points?	+1
Do the parties establish an SPS committee?	+1
Is there a fixed periodic meeting for the committee?	+1
Is the SPS committee the designated first place for dispute resolution?	+1
Does the SPS committee have open proceedings?	+1
Do the parties establish a working group?	+1
Is there a working group established?	+1
Is there a mechanism to issue recommendations?	+1
Is there a mechanism mandated to issue administrative decisions?	+1
Is a body for administering the agreement established?	+1
Is recourse to dispute settlement for the SPS chapter disallowed?	+1

Table 2: PTA design dimensions of interest in the measurement of DC institutionalization. Each indicator is weighted according to its theoretical relevance to the notion of deferred contracting (DC). These dimensions are retrieved from the World Bank Dataset on Deep Trade Agreements (Rocha et al., 2020).

For TBT and SPS, *Dc.Institutionalization* considers whether PTA clauses establish administrative bodies, if they institutionalize consultations to resolve disputes, and mechanisms for the parties to issue recommendations to one another, and if these recommendations are binding. Since DC is exercised through consultations

and bargaining, as opposed to adjudication, introducing a formal dispute settlement body counts negatively, while its disallowance scores positively. Indeed, adjudication typically entails delegation to third-party actors (Rosendorff, 2005), whereas DC is directly managed by the original negotiators of the agreement. States entering PTAs which score above the index mean are considered as treated, and untreated otherwise. I stress how limiting my focus to TBT and SPS also avoids issues arising from simply counting the occurrence of implementing institutions across all treaty chapters. The latter approach may lead to spurious results whereby the observed effect is determined by the sheer number of chapters of an agreement, as opposed to the degree of institutionalization of DC within each of them.

To ensure that the assumption of parallel trends is satisfied, I weight control dyads using propensity-score weighting on a series of covariates in the four years preceding treatment (t0, ...t-4). I weight on my outcome measure of regulatory distance; the log of the within-dyad gap in GDP; the gap in democracy levels and civil society involvement in the policy process, based on the Electoral Democracy and the Civil Society Participation indices from V-Dem Project (Coppedge et al., 2023); and on whether dyad members have raised specific trade concerns (STC) against each on TBT or SPS at the WTO. Weighting on these variables seeks to ensure that treatment and control units (country dyads) have highly similar histories in trade and regulation, distribution of bargaining power, and domestic institutions.

7 Empirical Findings

The analysis of my findings begins with Figure 7 (below), which displays results from the DiD estimation, which considers the sub-sample of country dyads having entered a PTA between 1995 and 2020 (n= 190). I plot the estimates of the average treated effects on the treated dyads (ATT) and 95% confidence intervals for the 6 years following PTA signature, which occurs at *t=0*. The post-treatment period corresponds with the unshaded area. I also report ATT estimates and 95% confidence intervals for placebo PTA signatures occurring between *t-4* and *t-1*, which correspond with the shaded area.

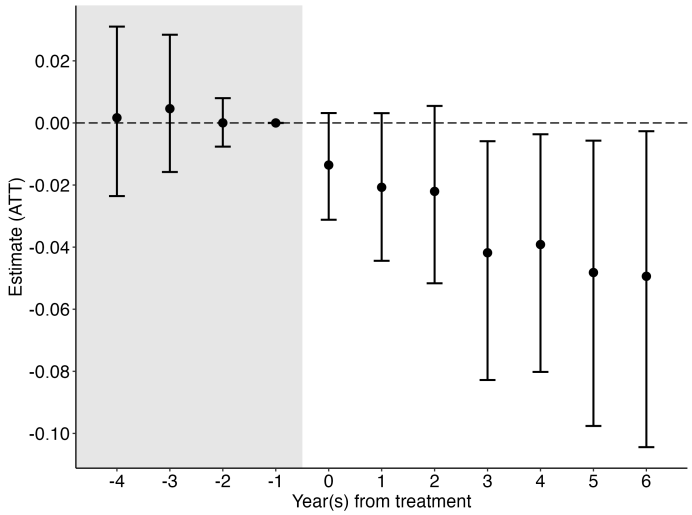


Figure 7: Effect of DC institutionalization on regulatory distance. The plot shows the ATT of entering a PTA with high DC institutionalization on the distance between the regulatory ideal points of its members. The model relies on propensity-score weighting and estimates 95% confidence intervals with 1000 bootstraps.

I find that entering a PTA establishing above-average DC institutions for TBT and SPS reduces the regulatory distance between treaty members by about 5% in the 6 years following signature. This effect is substantive considering that, as shown in Figure 4, the average regulatory worldwide has increased on average by around 1% over the past 3 decades. A PTA with high DC institutionalization can offset regulatory divergence between two countries by a magnitude that is about five times larger the average cross-country regulatory gap, relative to agreements lacking a similar institutional infrastructure. The point estimates indicate a clear and downward trend in regulatory distance up until the sixth year after signing. Interestingly, the downward trend acquires statistical significance at the 5% level only from *t+3* onward, that is 3 years after a PTA is concluded, and, on average, about 1.5 years since its ratification.[6] The shaded area of the plot displays placebo tests performed in the 4-year lag preceding treatment. As no ATT is statistically significant, it is possible to exclude that dyads experience pre-trends in regulatory convergence before the conclusion of

[6]The average time between PTA signature and ratification is 1.5 years.

a PTA and that the model satisfies the parallel-trend assumption of DiD. The presence of parallel trends is further validated by computing the balance between covariates used for ps-weighting, whose standardized differences are consistently around 0 in the 4-year lag before the treatment.[7]

The above findings shed important light on the effect DC has on bilateral regulatory cooperation in the aftermath of concluding a PTA. To start with, the conclusion of a PTA itself marks an important watershed moment in bilateral regulatory cooperation. While literature has generally argued that PTAs are endogenous to already intense trade relations (Baccini, 2019), my findings paint a different picture. They suggest that preferential trading partners do witness significant regulatory convergence, but do so only after the conclusion of PTAs and increasingly throughout the implementation years. I put particular emphasis on the theoretical importance of placebo tests in the above estimation. The lack of any significant pre-trend in convergence before a PTA is concluded suggests that international institutions still remain critical to non-tariff cooperation. They determine whether even closely integrated trade partners will embark on a process of aligning their product standardization systems — beyond making mutual tariff concessions. Indeed, while cooperation around tariffs has advanced ever since the establishment of the WTO, NTMs represent the principal obstacle to twenty-first century liberalization (Baldwin, 2011). At one level, NTMs present a hard test for international cooperation as their complexity and domestic nature makes them hard to govern. At another level, NTMs act as tariff substitutes that can serve protectionist interests where at-the-border measures no longer do so. In this light, the value of institutions like PTAs to regulatory cooperation emerges in their ability to structure long-term cooperation around complex policy problems and to build political support for behind-border liberalization.

Moving more closely to the implementation phase, the negative effect of DC institutionalization on regulatory distance becomes statistically significant only from the third year after the signing of a treaty. This is particularly telling of the mechanics of DC. While the literature has found that the majority of tariff lines are fully liberalized within the first year of implementation (Baccini et al., 2018), institutions of deferred contracting start impacting regulatory convergence only farther into implementation. As the qualitative case study on the EU Japan EPA further elaborates upon, this is not simply a function of the time required for translating more complex treaty commitments into concrete implementation measures. It is also determined by the timing of bargaining over additional mutual commitments and to demand domestic policy adjustments and concessions that are poorly legalized in the agreement itself.

[7] The covariate-balance plot is available in the Appendix, Table E.1.

8 Robustness Checks

I move to discussing a series of checks performed to assess the robustness of my main findings. I begin by considering the robustness of my outcome measure of regulatory distance to causal estimation. A different specification of the simple distance between the first principal components (PC1s) in a dyad is the Mahalanobis distance. Mahalanobis distance weighs raw distance measures based on the distribution of PC1 scores across dyads, penalizing large distance values and giving greater weights to small distance changes. I replicate my main estimation model by replacing the baseline regulatory distance measure with regulatory Mahalanobis distance. Figure 8 (below) plots the ATT estimates and 95% confidence intervals for the 6 years following PTA signature (*t=0,...,6*) as well as placebo tests (*t-1, ...,t-4*).

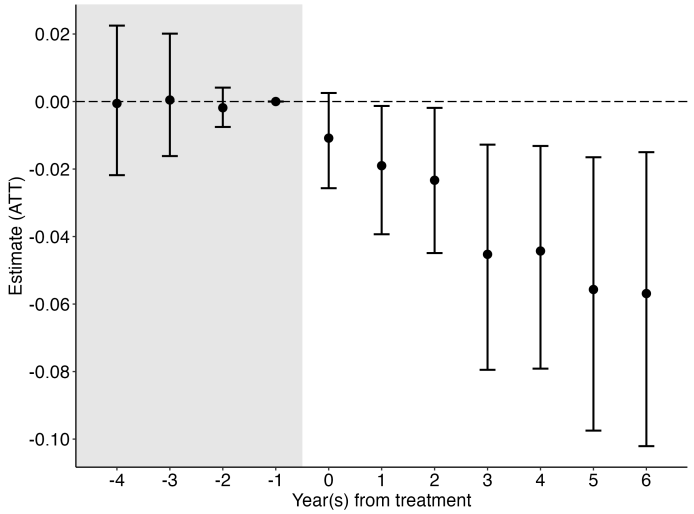
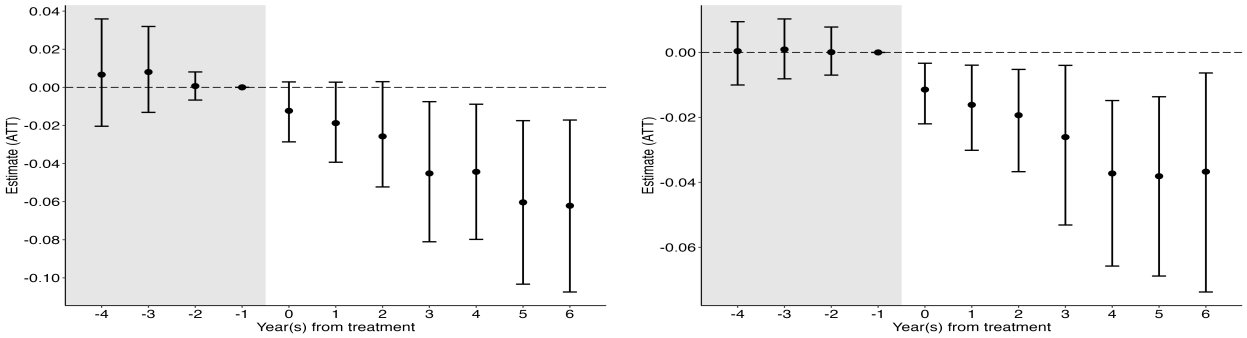


Figure 8: Effect of DC institutionalization on regulatory Mahalanobis distance. The plots show the ATT of entering a PTA with high DC institutionalization on the distance between the regulatory ideal points of its members. The models rely on propensity-score weighting and estimate 95% confidence intervals with 1000 bootstraps.

The ATT estimates above are highly consistent with the main results. They indicate that PTAs with more institutionalized DC lead their members to converge by up to 5% more in their regulations relative to signatories of other PTAs. These results are statistically significant at the 5% level. Additionally, the estimates for placebo tests (shaded region) are statistically insignificant in the entire pre-treatment period. This indicates that the assumption of parallel trends is met. Differently from my baseline estimation (Figure 7), ATT estimates for regulatory Mahalanobis distance indicate that convergence begins earlier after signature, compared to baseline results presented earlier. In particular, ATT estimates are negative and statistically significant starting from the first year following PTA signature. This does not weaken the principal feature of deferred contracting (DC) of enacting cooperation in the longer-run, rather than right upon entry into

force of the PTA: indeed, ATT estimates are twice as strong from *t+3* onwards, which still captures the long-term negative effect of DC dynamics on regulatory distance. [8]

To put baseline results into a wider perspective, it is necessary to also estimate the effect that entering a PTA has relative to no preferential cooperation between dyads. This requires expanding the sample of dyads to include states that never entered a PTA as the control group. I estimate the effect that entering a PTA establishing above/below-average DC institutions has on regulatory distance relative to states not bount by and agreement between 1995 and 2020. I run my estimation on two separate groups: (i) one including dyads entering PTAs with above-mean DC institutions and dyads forming no PTA; (ii) and one including dyads entering PTAs with above-mean DC institutions and dyads entering no PTA. These results are displayed in Figure 9 (below), which shows the ATT estimates for both models. The post-treatment period (*t=0,...,6*) and the placebo-test lag (*t=-1,...,-4*) are the same as in all previous models. For each sub-sample, the treatment variable scores 1 for country dyads signing a PTA in a given year, and 0 otherwise.[9]



(a) DC Institutionalization > Mean **(b)** DC Institutionalization < Mean

Figure 9: Effect of DC institutionalization on regulatory distance. The plots show the ATT of entering a PTA with DC institutionalization above the mean (plot a) and below the mean (plot b) on the distance between the regulatory ideal points of member dyads. The estimations rely on propensity-score weighting and estimate 95% confidence intervals with 1000 bootstraps.

Results in Figure 9 are consistent with my hypothesis that a higher number of institutions for DC entails stronger regulatory convergence (or reduction of regulatory distance) between treaty members. ATT estimates suggest that more institutionalized DC leads country dyads to converge by up to 8% in the six years following signature, relative to dyads not bound by a PTA. Coherently with my theory that DC yields its effects mainly after ratification, ATT estimates in plot (a) only become statistically significant from the third year after signature onward. Plot (b) suggests that PTAs with low DC institutionalization only reduce regulatory distance by up to 4%, and do so significantly right from the year of signature (*t0*). Interestingly,

[8] The covariate-balance plots are available in the Appendix, Table E.2.
[9] The covariate-balance plots are available in the Appendix, Tables E.3.

agreements lacking a high number of DC institutions seem to enact regulatory convergence more quickly, but to a lesser degree. This suggests that formal treaty clauses are enough to translate shallower commitments to harmonization into action. In contrast, deeper harmonization goals require significant heavy lifting ex-post, which is aided by the presence of DC fora.

9 Case Evidence: the EU-Japan EPA

The conclusion of the EPA, which was ratified in February 2019, was in line with Japan's policy of wide support for preferential market integration, including through greater agricultural liberalization. From an EU side, the EPA represented the culmination of over a decade of crafting new-generation PTAs, since the launching of its "Global Europe" Strategy, which aimed at going beyond commitments in the WTO and pursuing the diffusion of Europe's regulatory standards globally (European Commission, 2006). Abenomics' focus on free trade had to be reconciled with Japan's traditional attention to protecting the politically sensitive "five sacred items:" rice, wheat, pork and beef, dairy, and sugar (Davis, 2022; Maclachlan & Shimizu, 2022). This was a tricky equilibrium to maintain considering that the EPA brought ambitious agricultural liberalization. Figures from the agreement indicate a progressive elimination of tariffs across 97% of tariff lines. This includes sensitive agricultural products. For example, EU hard cheese imports are set to be fully liberalized by 2033, beef imports were set to undergo a 30% tariff cut over 15 years, while wine underwent an immediate tariff drop from 15% to 0% as the EPA came into force (European Commission, 2018a).

These costly commitments were also made possible by issue-linking Japan's agricultural liberalization to the EU's commitment to lift its 10% import tariff on cars (Suzuki, 2017). Additionally, EU concessions on cars were welcomed by the powerful industrial lobbying group *Keidanren* and by corporate giants like Toyota. In contrast, a principal objective of the powerful farmer lobby Japan Agricultural Cooperatives (JA), or *Nōkyō*, was postponing as much as possible the envisioned agricultural liberalization. A 2017 JA position paper on the EPA already signaled that EU demands for Japan's reforms of its SPS measures on cheese, pork, and wine would be a major source of contention under the agreement (Central Union of Agricultural Cooperatives, 2017). Indeed, since the coming into force of the EPA, the power to delay a surge in EU agricultural imports shifted from tariffs to non-tariff measures, especially sanitary and phytosanitary (SPS) measures (Suzuki, 2017; Vorst, 2021).

9.1 Deferred Contracting in the EPA

The EPA extensively deferred contracting on SPS to the implementation phase, by establishing an SPS Joint Committee endowed with decisional powers extending beyond the treaty text. EPA Article 6.15 specifies that the SPS Committee 'may identify and consider technical cooperation projects between the Parties in relation to the development, implementation, and application of sanitary and phytosanitary measures', while also acting as a forum for the EU and Japan to coordinate their policies at WTO level. The Joint Committee gathers representatives from Japan's Ministry of Foreign Affairs (MOFA), the Ministry of Agriculture, Forestry and Fisheries (MAFF), and the Ministry of Health, Labour and Welfare (MHLW). On the EU side, it includes representatives from the Commissions Directorates General for Trade, Health, as well as national delegates from the rotating Council presidency. At the time of writing, the SPS Joint Committee has met five times (once a year), with national bureaucracies coordinating their policies regularly in preparation for those meetings.

The conclusion of the EPA itself enhanced significant SPS convergence between the EU and Japan. Yet, Japan's regulatory framework for SPS has remained considerably more restrictive than the EU's ever after the agreement was ratified. Indeed, the EPA granted Japan considerable discretion in managing agricultural imports from different regions of the EU, evaluating their approval on a case-by-case basis. Japan has maintained lengthy SPS approval procedures for various tariff schedules, notably dairies and beef. Most importantly, Japan did not streamline these verification requirements, replicating them on an individual basis for each applying Member State in the EU. This was institutionally allowed by the loose phrasing of the EPA Article 6.7.2, which states that 'the importing Party *should*, if requested by the exporting Party, apply the import conditions for products to the entire territory of the exporting Party in a consistent manner' (emphasis added). If an EU Member State undergoes an approval procedure for a product regulated under a given SPS rule in the EU, an identical the request by another Member State is treated by Japan separately.

In the first meeting of the SPS Committee, held in October 2019, EU officials lamented Japan's lengthy SPS import and risk-assessment procedures, adding them to the agenda as a general SPS concern. The issue was raised consistently during subsequent meetings. In the agenda of the Committee's second meeting, the EU and Japan officially tabled the 'simplification of the Japanese SPS procedures for import and the risk assessment of EU Member States' (European Commission, 2021b, 13). Considering that Japan has no binding obligation to harmonize risk assessment procedures towards EU Member States (as per Art. 6.7.2), these efforts went beyond formal commitments in the treaty. The EU has pushed for a harmonized approach due to the import bottleneck that country-by-country assessments create, while Japan is yet to fully streamline these procedures, at the time of writing. In particular, during the third SPS Joint Committee

meeting, held in December 2021, the EU proposed to set up meetings between technical experts on the issue of Japan's differential treatment and committed to share comments and questions in writing to define items to be discussed during those meetings (European Commission, 2021a).

Japan's approach to risk-assessment reveals a protectionist logic. Given formal liberalization commitments on SPS in the EPA, a country-by-country approach to risk assessment seems costly and counterintuitive: it cannot systematically block the inflow of EU agricultural products. It can only delay it, while multiplying costly verifications. Similar delaying strategies were a key ask from *Nōkyō* and the MAFF. While the MAFF showed broad support for Abe's reforms, including greater centralization of trade policy around the Prime Minister's Office (*Kantei*), it was consistent in arguing that agricultural liberalization should be strictly limited or postponed (Suzuki, 2017, 882). In sum, agricultural liberalization was still a costly concession for Japan as it entered the EPA.

Desptite agriculture's sensitivity for Japan, the bargaining setting of the SPS Joint Committee has granted MAFF officials considerable policy discretion away from key constituencies. First, administrative decisions on SPS verifications do not require approval by the National Diet. Second, the participation of the MAFF in the EPA could rely on a new generation of more pro-trade bureaucrats developing within the Ministry (Maclachlan & Shimizu, 2022). These agents met regularly with their EU counterparts both at Joint Committee meetings and as part of the regular policy coordination in preparation for those meetings. The EU's 2021 EPA Implementation Report already recorded progress in Japan's risk assessment procedures for certain EU beef products, two years into implementation, which suggests Japan has progressively yielded to pressures from the EU (European Commission, 2021b). The latest implementation report, issued in 2024, also recorded full convergence on checks related to avian influenza for poultry (European Commission, 2024). Throughout the EPA's implementation, Japan rolled back its practice of applying EU-wide import bans on poultry even in instances of highly localized influenza cases.

Pointing to the fundamentally contractual nature of these exchanges in the SPS Committee is the fact that Japan's further liberalization was linked to additional concessions from the EU. Official minutes indicate that up until the end of 2021 (three years into implementation), EU officials were urged by their Japanese counterparts to lift the import ban on agricultural products coming from the area of the Fukushima nuclear power plant, which witnessed radioactive spills in 2011. Shortly after ratification, the EU sent a strong signal of its intention to lift the ban by passing an amending regulation (Regulation (EU) 2019/1787). Yet, a formal commitment to lifting this ban was never included into the EPA text itself, due to the political sensitivity the issue had amid the European public. This suggests how DC dynamics can routinely take bilateral policy coordination beyond just passive implementation. After the passing of the amending regulation, Japan brought up the issues consistently at Joint Committee meetings until 2021, urging the EU to enforce its

regulatory change, which the EU fully enacted only in 2023 (European Commission, 2023).[10]

A structural advantage of the EPA network of specialized committees is the fact that these bodies lacked exposure to Japan's agricultural lobbies. While representatives of the MAFF sit in the SPS Committee, no structured representation of the *Nōkyō* exists. In fact, the main business group overlooking EPA implementation talks has been the EU-Japan Business Round Table, a conglomerate of firms and trade associations which has surprisingly no representation from the agricultural sector, but gathers, on the Japanese side, 33 members from the tech and transportation sector. The Round Table has tended to focus more on issues relevant to industrial manufacturing, like TBTs, for instance by pressuring the Japanese government to converge towards UNECE car safety standards already in force in the EU (Vorst, 2021). Institutionally, the issue-specific compartmentalization of the the Joint Committee also put considerable pressure on Japan to ease its SPS barriers. The topical focus of the SPS Committee leaves little room for wide-scope issue-linkage, making it harder for Japan to resist on agricultural matters while making concessions in other, less costly areas.

Overall, official SPS Joint Committee reports indicate that the issue of Japan SPS risk assessments and the EU lifting of the Fukushima ban, while being minimally or not addressed in the EPA itself, were core agenda items throughout the bilateral talks that followed the ratification of the agreement. They were discussed by the same delegations, in the same rooms, where each meeting round required several months of preparation. Crucially, official documents reveal that the two issues were concomitantly on the agenda for at least three years after EPA enforcement. It would be speculative, in the absence of more data, to argue that the linkage between these two issues was actively leveraged by either party to achieve a breakthrough in reciprocal SPS liberalization. Yet, it is documented that both the streamlining of Japan's risk assessment procedures and the EU lifting of the Fukushima ban were repeatedly discussed after treaty enforcement by the SPS Joint Committee, and, as in the case of the EU ban, also seemingly resolved. The policy implications of these efforts are substantive, considering that the EU and Japan form the largest bilateral trade partnership worldwide and that Japan is the fifth largest recipient of total EU agricultural export.

10 Alternative Explanations: What about the WTO?

For the 166 member states of the WTO, any PTA they conclude is embedded in the system of multilateral rules set by the Organization. The WTO Agreements on TBT and SPS constitute the main multilateral framework governing these fields. These treaties established two governing Committees for TBT and SPS

[10]The central importance of lifting the Fukushima ban for Japan should not be underestimated, given that failure to agree on this concession ended PTA negotiations with other trade partners like Taiwan and has delayed Taiwan's entry into the CPTPP (Hou & Ellis, 2022; Strong, 2019).

issues, where members can raise specific trade concerns (STCs) during the regular meetings of the TBT and SPS Committees. STCs can be issued in response to government notifications of newly introduced regulations, but also target measures having been in place for some time (Holzer, 2018). STCs represent an important alternative institution explaining why two countries tied by a PTA and also members of the WTO would converge in their regulatory practices.

Similarly to policy dialogues within PTA joint committees, STCs are also regularized in that they can be raised at the multiple yearly meetings held by the TBT and SPS Committees. The agenda of these meetings are also specific as STCs usually revolve around individual measures adopted by "concerning" states over specific product categories. The STC procedure is closely modeled upon the more formal disputes submitted to the Dispute Settlement Mechanism (DSM). In this light, STCs have been regarded as a cost-efficient device favoring the diplomatic resolution of regulatory controversies before these would escalate into more formal and costlier dispute-settlement tools (Holzer, 2018).

However, the functioning of STCs implies considerable selection of the regulatory concerns that are eventually raised. STCs focus on contentious items reaching a certain salience threshold for domestic businesses and constituents that does not allow the issue to be settled by more informal means. The literature has shown how domestic business groups lobby their governments on what disputes to initiate, while overlooking many other potential ones (Brutger, 2023; Davis, 2012). It follows that, however frequent, the use of STCs constitutes a limited approach to regulatory convergence. Irrespective of the successful resolution of a trade concern, the impact of an STC on aggregate levels of regulatory convergence between two member states can be expected to be marginal.

I test the effect that raising an STC has on the regulatory distance for dyads of WTO members. I rely on the same DiD estimator used in previous models. The unit of analysis is the country dyad of WTO members, combined pair-wise using all possible combinations in the period ranging between 1995 and 2023 included. The treatment variable scores 1 for country dyads initiating an STC in a given year, and 0 otherwise. I weight country pairs on the same range of covariates used in earlier models, controlling for whether the dyad members have entered a PTA in each of the 4 years prior to treatment. Figure 13 (below) plots the average treatment effect on the treated (ATT) for the effect of raising an STC on the regulatory distance of WTO member dyads. Dyad members having initiated an STC against one another in a given year are considered treated, and untreated otherwise.

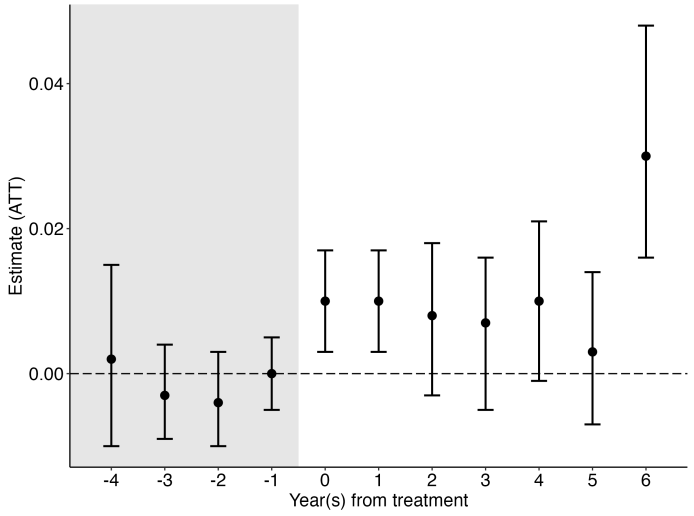


Figure 10: ATT estimates for the effect of raising an STC at the WTO (unshaded area). Placebo estimates with 95% block-bootstrap confidence intervals (shaded area). The model rely on propensity-score weighting and estimates 95% confidence intervals with 1000 bootstraps.

The results in Figure 10 (above) do not show support for the alternative explanation that raising STCs at the WTO significantly reduces regulatory distance between two countries. The post-treatment (*t=0,1,...,6*) and the placebo test periods (*t=-4,...,-1*) are the same as in earlier models. ATT point estimates are both positive and close to zero, but their 95% confidence intervals mostly indicate a lack of statistical significance in the post treatment period, except for *t0*, *t+1*, *t+6*, where the estimates are surprisingly positive. It can be cautiously concluded that raising an STC, while lacking any consistent effect on regulatory distance, denotes that two WTO disputing members are likely embarking on a regulatory-divergence path that is not offset, on average, by WTO TBT and SPS fora.

11 Conclusion

The institutionalization of deferred contracting (DC) in international agreements enhances cooperation on costly commitments during the implementation of international agreements. After the ratification of a treaty, government officials can participate in regularized meetings, joint committees, and exercise binding ex-post decisional powers to make mutual concessions that are not strictly formalized in the original contract. In the field of TBT and SPS, higher institutionalization of DC has led to greater regulatory convergence between PTA signatories. It has on average reduced the regulatory distance for 283 country dyads by up to 5% in the course of the six years following signature. Countries concluding PTAs producing more institutions for DC converge more than other countries.

Resorting to DC can lead to important liberalization breakthroughs on costly commitments that are

difficult to formalize during the negotiations of the original contract. DC institutions do not simply enforce. They are able to make policy decisions that are not merely derivative of formal treaty commitments, but serve to expand, modify, or retrench them. This can be achieved as policy decision-making is shielded from the pluralistic constituency oversight typical of open negotiations, where a variety of stakeholders ranging from citizen movements to firms can mobilize around specific focal point to jeopardize treaty ratification. DC is also sustained by domestic businesses carrying sunk costs from their lobbying in favor of PTA adoption, and demanding that implementation delivers on the treaty commitments they have been pushing for.

Qualitative case evidence is also coherent with statistical findings in this paper. Investigating SPS cooperation following the ratification of the EU–Japan EPA, I have argued that DC institutions enable treaty members to cooperate on issues that are scantly legalized in treaty texts and which, if not addressed ex post, can significantly limit liberalization. Being able to regularly perform these functions is especially key in the field of regulatory cooperation. In international trade, the rise of non-tariff barriers over the last three decades suggests that a large share of modern liberalization is about navigating the intricacy of regulatory barriers, that is technical policies which, unlike tariffs, require institutionalized transparency and time to become visible and governable (Kono, 2006).

While open negotiations represent the phase at which political commitments to liberalize are discussed and openly contested, implementation is the phase where many of the terms of cooperation are crafted and set in motion. In this light, the politics of treaty implementation can determine considerable differences in the economic effects and efficiency gains determined by formally similar international agreements. The power of DC to shape the terms of cooperation increases in the incompleteness of signed contracts. This is especially relevant to the trade regime governing non-tariff measures, where both the GATT/WTO and PTAs have failed to produce clear and binding provisions. Overall, although it is unlikely that implementation per se entirely sets the terms of given international policies, understanding how this overlooked phase governs weakly legalized and salient treaty commitments can improve our understanding of cooperation under international institutions.

References

Abbott, K. W., Keohane, R. O., Moravcsik, A., Slaughter, A. M., & Snidal, D. (2000). The Concept of Legalization [Publisher: Cambridge University Press]. *International Organization*, *54*(3), 401–419. https://doi.org/10.1162/002081800551271

Abbott, K. W., & Snidal, D. (2000). Hard and Soft Law in International Governance [Publisher: Cambridge University Press]. *International Organization*, *54*(3), 421–456. https://doi.org/10.1162/002081800551280

Baccini, L. (2019). The Economics and Politics of Preferential Trade Agreements [Publisher: Annual Reviews ISBN: 2019.22:7592]. *Annual Review of Political Science*, *22*, 75–92. https://doi.org/10.1146/ANNUREV-POLISCI-050317-070708

Baccini, L., Dür, A., & Elsig, M. (2018). Intra-Industry Trade, Global Value Chains, and Preferential Tariff Liberalization [Publisher: Oxford Academic]. *International Studies Quarterly*, *62*(2), 329–340. https://doi.org/10.1093/ISQ/SQY006

Baldwin, R. E. (2011). 21st Century Regionalism: Filling the Gap between 21st Century Trade and 20th Century Trade Rules. https://doi.org/10.2139/ssrn.1869845

Barton, J., Goldstein, J., Josling, T., & Steinberg, R. (2008). *The Evolution of the Trade Regime — Princeton University Press* [ISBN: 9780691136165]. Princeton University Press. Retrieved August 18, 2025, from https://press.princeton.edu/books/paperback/9780691136165/the-evolution-of-the-trade-regime

Berggren, N., Bergh, A., & Bjørnskov, C. (2012). The growth effects of institutional instability. *Journal of Institutional Economics*, *8*(2), 187–224. https://doi.org/10.1017/S1744137411000488

Biffi, L. (2020). *Too Complex or too Ambitious? Exploring hurdles in the enforcement of EU free trade agreements* [Doctoral dissertation, The Graduate Institute of International Studies and Development, Geneva].

Bögner, A. (2024). Trading off climate: How conventional trade interests shape climate discussions at PTA committees [Publisher: Routledge]. *Review of International Political Economy*. https://doi.org/https://doi.org/10.1080/09692290.2024.2428782

Bombardini, M. (2008). Firm heterogeneity and lobby participation [Number: 2]. *Journal of International Economics*, *75*(2), 329–348. https://doi.org/10.1016/j.jinteco.2008.03.003

Brutger, R. (2023). Litigation for Sale: Private Firms and WTO Dispute Escalation. *American Political Science Review*, 1–18. https://doi.org/10.1017/S0003055423000850

Central Union of Agricultural Cooperatives. (2017). *Concerns Regarding Agricultural Trade in the EU–Japan EPA, EU* (tech. rep.). Tokyo. https://www.zenchu-ja.or.jp/wp_zenchu/wp-content/uploads/2022/09/up186.pdf

Claussen, K. (2024). Operationalizing Package Treaties: A US Case Study. *World Trade Review*, *23*(5), 589–600. https://doi.org/10.1017/S1474745624000478

Coppedge et al. (2023). V-Dem Codebook v13. https://www.v-dem.net/documents/24/codebook_v13.pdf

Czechowska, L. (2022). Joint Bodies and the Regularization of Strategic Interaction: A Comparison of the European Union's Strategic Partnerships with Japan and India [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111 *JCMS: Journal of Common Market Studies*, *60*(4), 1144–1164. https://doi.org/10.1111/jcms.13325

Davis, C. L. (2004). International Institutions and Issue Linkage: Building Support for Agricultural Trade Liberalization [Publisher: Cambridge University Press]. *American Political Science Review*, *98*(1), 153–169. https://doi.org/10.1017/S0003055404001066

Davis, C. L. (2012). *Why Adjudicate? — Princeton University Press* [ISBN: 9780691152769]. Princeton University Press. Retrieved March 17, 2024, from https://press.princeton.edu/books/paperback/9780691152769/why-adjudicate

Davis, C. L. (2022). Japanese Trade Policy. In R. J. Pekkanen & S. M. Pekkanen (Eds.), *The Oxford Handbook of Japanese Politics* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190050993.013.30

Department of the Treasury. (2025). Treasury Releases Proposed Guidance to Continue U.S. Manufacturing Boom in Batteries and Clean Vehicles, Strengthen Energy Security. Retrieved March 26, 2025, from https://home.treasury.gov/news/press-releases/jy1939

Downs, G. W., Rocke, D. M., & Barsoom, P. N. (1996). Is the good news about compliance good news about cooperation? *International Organization*, *50*(3), 379–406. https://doi.org/10.1017/S0020818300033427

Doyle, D., & Elgie, R. (2016). Maximizing the Reliability of Cross-National Measures of Presidential Power. *British Journal of Political Science*, *46*(4), 731–741. https://doi.org/10.1017/S0007123414000465

Dür, A., & Gastinger, M. (2023). Spinning a global web of EU external relations: How the EU establishes stronger joint bodies where they matter most [Publisher: Routledge _eprint: https://doi.org/10.1080/13501763.2022.20 *Journal of European Public Policy*, *30*(6), 1072–1091. https://doi.org/10.1080/13501763.2022.2079708

Dür, A., & Mateo, G. (Eds.). (2016). *Insiders versus Outsiders: Interest Group Politics in Multilevel Europe*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198785651.002.0003

Espitia, A., Pardo, S., Piermartini, R., & Rocha, N. (2020). Technical Barriers to Trade. In A. Mattoo, N. Rocha, & M. Ruta (Eds.), *Handbook of Deep Trade Agreements* (pp. 343–366). World Bank Group.

https://documents1.worldbank.org/curated/en/685311594363725995/pdf/Handbook-of-Deep-Trade-Agreements.pdf

European Commission. (2006). Global Europe: Competing in the World. A Contribution to EU's Growth and Jobs Strategy. COMM(2006) 567 [Pages: 1-18]. Retrieved June 20, 2020, from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2006:0567:FIN:en:PDF

European Commission. (2018a). Key elements of the EU-Japan Economic Partnership Agreement. Retrieved April 6, 2025, from https://ec.europa.eu/commission/presscorner/detail/en/memo_18_6784

European Commission. (2018b). Meeting of First Sanitary and Phytosanitary Joint Management Committee (JMC) [Pages: 1-2 Place: Ottawa, 26-27 March 2018].

European Commission. (2021a). *3rd meeting of the Committee on Sanitary and Phytosanitary Measures (7-9 December 2021)* (tech. rep.). https://www.mofa.go.jp/mofaj/files/100297037.pdf

European Commission. (2021b). *EU-Japan EPA Progress Report* (tech. rep.). https://op.europa.eu/en/publication-detail/-/publication/6545cd8a-5912-11ed-92ed-01aa75ed71a1/language-en

European Commission. (2023). Lifting of Fukushima restrictions on food imports. Retrieved April 8, 2025, from https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3781

European Commission. (2024). *Report from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on the implementation and enforcement of EU trade policy* (tech. rep.). https://ec.europa.eu/commission/presscorner/detail/en/ip_24_5021

Garcés, I., & Vogt, A. (2024). Patterns of regulatory heterogeneity in international trade: Intensity, coverage, and structure [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/roie.12736]. *Review of International Economics*, *n/a*(n/a). https://doi.org/10.1111/roie.12736

Goldstein, J., & Martin, L. L. (2000). Legalization, Trade Liberalization, and Domestic Politics: A Cautionary Note. *International Organization*, *54*(3), 603–632. https://doi.org/10.1162/002081800551226

Gulotty, R. (2020). *Narrowing the Channel: The Politics of Regulatory Protection in International Trade*. The University of Chicago Press.

Hart, O., & Moore, J. (1988). Incomplete Contracts and Renegotiation [Publisher: [Wiley, Econometric Society]]. *Econometrica*, *56*(4), 755–785. https://doi.org/10.2307/1912698

Holzer, K. (2018, October). *Addressing Tensions and Avoiding Disputes: Specific Trade Concerns in the TBT Committee* (tech. rep.). WTO Economic Research and Statistics Division.

Hou, B., & Ellis, S. (2022). Taiwan Lifts Ban on Fukushima Food in Push to Join Trade Bloc. *Bloomberg.com*. Retrieved April 6, 2025, from https://www.bloomberg.com/news/articles/2022-02-07/taiwan-set-to-lift-ban-on-fukushima-food-to-boost-japan-ties

Imai, K., Kim, I. S., & Wang, E. H. (2023). Matching Methods for Causal Inference with Time-Series Cross-Sectional Data [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12685]. *American Journal of Political Science*, *67*(3), 587–605. https://doi.org/10.1111/ajps.12685

Jackman, R. W., & Miller, R. A. (1996). A Renaissance of Political Culture? [Publisher: [Midwest Political Science Association, Wiley]]. *American Journal of Political Science*, *40*(3), 632–659. https://doi.org/10.2307/2111787

Kim, I. S. (2018). *LobbyView: Firm-level Lobbying & Congressional Bills Database*. https://web.mit.edu/insong/www/pdf/lobbyview.pdf

Kono, D. Y. (2006). Optimal Obfuscation: Democracy and Trade Policy Transparency. *American Political Science Review*, *100*(3), 369–384. https://doi.org/10.1017/S0003055406062241

Koremenos, B. (2005). Contracting around International Uncertainty. *American Political Science Review*, *99*(4), 549–565. https://doi.org/10.1017/S0003055405051877

Koremenos, B., Lipson, C., & Snidal, D. (2001). The Rational Design of International Institutions. *International Organization*, *55*(4), 761–799. https://doi.org/10.1162/002081801317193592

Laurens, N. D., Hollway, J., & Morin, J. F. (2023). Checking for Updates: Ratification, Design, and Institutional Adaptation [Publisher: Oxford Academic]. *International Studies Quarterly*, *67*(3), 49. https://doi.org/10.1093/ISQ/SQAD049

Lavenex, S. (2009). Switzerland's Flexible Integration in the EU: A Conceptual Framework [_eprint: https://onlinelibrary.wiley.
6370.2009.tb00145.x]. *Swiss Political Science Review*, *15*(4), 547–575. https://doi.org/10.1002/j.1662-6370.2009.tb00145.x

Maclachlan, P. L., & Shimizu, K. (2022). The Farm Lobby and Agricultural Policy in Japan. In R. J. Pekkanen & S. M. Pekkanen (Eds.), *The Oxford Handbook of Japanese Politics* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190050993.013.22

Maggi, G. (2011). *Trade Agreements as Incomplete Contracts* (tech. rep.). National Bureau of Economic Research (NBER).

Maggi, G., & Ossa, R. (2020). Are Trade Agreements Good For You? *NBER Working Paper 27252*, 1–49.

Maggi, G., & Ossa, R. (2021). The Political Economy of Deep Integration [Publisher: Annual Reviews]. *Annual Review of Economics*, *13*, 19–38. https://doi.org/10.1146/ANNUREV-ECONOMICS-121020-032425

Magyar, Z. B. (2022). What Makes Party Systems Different? A Principal Component Analysis of 17 Advanced Democracies 1970–2013. *Political Analysis*, *30*(2), 250–268. https://doi.org/10.1017/pan.2021.21

Majone, G. (1994). The rise of the regulatory state in Europe. *West European Politics*, *17*(3), 77–101. https://doi.org/10.1080/01402389408425031

Morin, J.-F., Brandi, C., & Schwab, J. (2024). Environmental agreements as clubs: Evidence from a new dataset of trade provisions. *The Review of International Organizations*, *19*(1), 33–62. https://doi.org/10.1007/s11558-023-09495-3

PAN Europe. (2015). EU health policy on endocrine disruption collatoral damage in Commission health service SANTE's power play. Retrieved June 14, 2020, from https://www.pan-europe.info/press-releases/2015/05/eu-health-policy-endocrine-disruption-collatoral-damage-commission-health

Perlman, R. L. (2023). *Regulating Risk*. Cambridge University Press.

Putnam, R. D. (1988). Diplomacy and domestic politics: The logic of two-level games. *International Organization*, *42*(3), 427–460. https://doi.org/10.1017/S0020818300027697

Rocha, N., Ruta, M., & Fernandes, A. (2021). The economics of deep trade agreements: A new eBook. *VoxEU Column: International Trade*.

Rocha, N., Ruta, M., & Mattoo, A. (2020). *Publication: Handbook of Deep Trade Agreements*. World Bank Group. http://hdl.handle.net/10986/34055

Rogers, S., & Girolami, M. (2011). *A First Course in Machine Learning* [Google-Books-ID: rdQ1daD8BH8C]. CRC Press.

Rosendorff, B. P. (2005). Stability and Rigidity: Politics and Design of the WTO's Dispute Settlement Procedure [Publisher: Cambridge University Press]. *American Political Science Review*, *99*(3), 389–400. https://doi.org/10.1017/S0003055405051737

Rosendorff, B. P., & Milner, H. V. (2001). The Optimal Design of International Trade Institutions: Uncertainty and Escape [Publisher: Cambridge University Press]. *International Organization*, *55*(4), 829–857. https://doi.org/10.1162/002081801317193619

Shang, L., Lin, J.-C., & Saffar, W. (2021). Does economic policy uncertainty drive the initiation of corporate lobbying? *Journal of Corporate Finance*, *70*, 102053. https://doi.org/10.1016/j.jcorpfin.2021.102053

Stone, S., & Casalini, F. (2020). Sanitary and Phytosanitary Measures. In A. Mattoo, N. Rocha, & M. Ruta (Eds.), *Handbook of Deep Trade Agreements* (pp. 367–390). World Bank Group. https://documents1.worldbank.org/curated/en/685311594363725995/pdf/Handbook-of-Deep-Trade-Agreements.pdf

Strong, M. (2019). Taiwan missed a chance to sign FTA due to Fukushima food ban, May 21, 2019. Retrieved April 6, 2025, from https://taiwannews.com.tw/en/news/3707529

Suzuki, H. (2017). The new politics of trade: EU-Japan. *Journal of European Integration*, *39*(7), 875–889. https://doi.org/10.1080/07036337.2017.1371709

Tsebelis, G. (1995). Decision Making in Political Systems: Veto Players in Presidentialism, Parliamentarism, Multicameralism and Multipartyism. *Journal of Political Science*, *25*(3), 289–325. Retrieved May 25, 2020, from https://www.jstor.org/stable/194257

Tsebelis, G. (2011). *Veto players: How political institutions work* [ISSN: 02105233]. Princeton Paperbacks. https://doi.org/10.2307/40184562

UNCTAD. (2019). International Classification of Non-Tariff Measures. https://unctad.org/system/files/official-document/ditctab2019d5_en.pdf

UNCTAD. (2024). TRAINS Portal. https://trainsonline.unctad.org/home

Vaubel, R. (2006). Principal-agent problems in international organizations. *The Review of International Organizations*, *1*(2), 125–138. https://doi.org/10.1007/s11558-006-8340-z

Vorst, C. V. d. (2021). The non-tariff barriers of the EU–Japan EPA [Section: Japan, the European Union and Global Governance]. In *Japan, the European Union and Global Governance* (pp. 87–111). Edward Elgar Publishing. Retrieved May 22, 2024, from https://www.elgaronline.com/edcollchap/edcoll/9781788114455/9781788114455.00012.xml

You, H. Y. (2017). Ex Post Lobbying [Publisher: The University of Chicago Press]. *The Journal of Politics*, *79*(4), 1162–1176. https://doi.org/10.1086/692473

Young, A. R. (2016). Not your parents' trade politics: The Transatlantic Trade and Investment Partnership negotiations. *Review of International Political Economy*, *23*(3), 345–78. https://doi.org/10.1080/09692290.2016.1150316

A Principal-Component Analysis (PCA): Intuition

For each year, countries are observed as a "cloud" of datapoints distributed in space based on their frequency of regulatory measures cumulatively adopted under the 96 product groups of interest, i.e. the total number of product families in the Harmonized System (HS). To summarize information about each country's regulatory profile in this 96-dimensional space, PCA orthogonally projects each of the datapoints (countries) on the axis that captures their maximum variance, that is the first principal component (PC1). Since the data matrix is standardized, each PC1 score expresses how much a country deviates from the average pattern of regulation, with the sample mean of PC1 being equal to 0. The projection of the datapoints onto the one-dimensional space of the PC1 is visually summarized in Figure 5 (below).

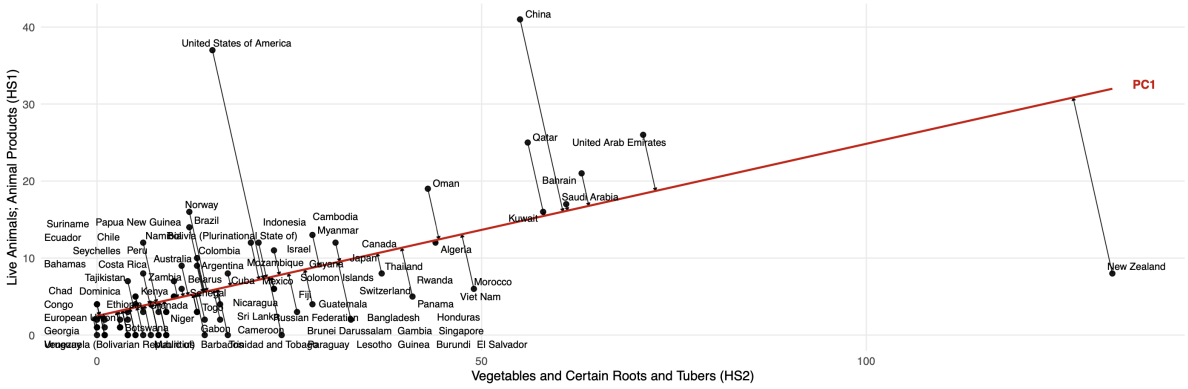


Figure 11: Intuition behind PC1 given a dataset containing countries' regulatory densities for two product families (HS1, HS2) in the year 2020. For visualization purposes, this plots displays the PCA shrinkage of 2 original dimensions, although this technique can handle many more dimensions, as in the case of the 96 original dimensions in the TRAINS data.

The projection of datapoints onto the red line (PC1) can be understood as a linear combination of the original dimensions allowing to project the original data-matrix of $D = 96$ dimensions onto a subspace of dimensionality $M < D$. The projection of the original two-dimensional data Y_n onto the one-dimensional X_n, as summarized in Figure 1, can be expressed through the following vector notation (Rogers & Girolami, 2011):

$$x_{nm} = W_m^{\mathrm{T}} Y_n \tag{1}$$

where x_{nm} represents each element of the m-th projection onto the principal subspace M (PC1), Y_n is the original column vector of D dimensions, and W_m^{T} is the transposed m-th eigenvector projecting the original datapoints in Y_n onto the M-dimensional subspace. PC1 coincides with the M-dimensional subspace where data variance is maximized, making the amount of variance the guiding criterion of the learning task for choosing W_m. Restricting this intuition to the 96 dimensions of interest (i.e., the 2-digit HS product families) in the regulatory-density data, equation 1 is specified as follows:

$$x_{nm} = W_m^{\mathrm{T}} Y_n, \quad where\ n = 96\ and\ m = 1 \tag{2}$$

PC1 is obtained by maximizing the variance of the projected data, which is given by the following equation:

$$\sigma_x^2 = \frac{1}{N}\sum_{n=1}^{N} x_n^2 = \frac{1}{N}\sum_{n=1}^{N} \left(\mathbf{w}^{\mathrm{T}}\mathbf{y}_n\right)^2 = \mathbf{w}^{\mathrm{T}} C\, \mathbf{w}. \tag{3}$$

Here, C is the covariance matrix of the (mean-centered) data $\{\mathbf{y}_n\}$. PCA constrains $\mathbf{w}$ to be a unit vector so that

$$\mathbf{w}^{\mathrm{T}}\mathbf{w} = 1. \tag{4}$$

Hence, given an eigenvalue/eigenvector pair $(\lambda, \mathbf{w})$ of C, the projection that maximizes the variance corresponds to the pair with the largest eigenvalue λ_1; in that case the maximal variance is $\sigma_x^2 = \lambda_1$ and the direction is $\mathbf{w}_1$.

B PC1 and PC2: Diagnostics

Figure 12 below displays the share of data variance explained by the first principal component (PC1) and other principal components up to PC10. PC1 alone accounts for around 50% of the total variance in the data and PC2 for just above 10%. PC3 and PC4 explain less than 10% of data variance while the remaining principal components up to PC10 score around 5%. It is therefore feasible to only discuss PC1 and PC2 for the remaining of this analysis, as the two together account for about 60% of overall data variance.

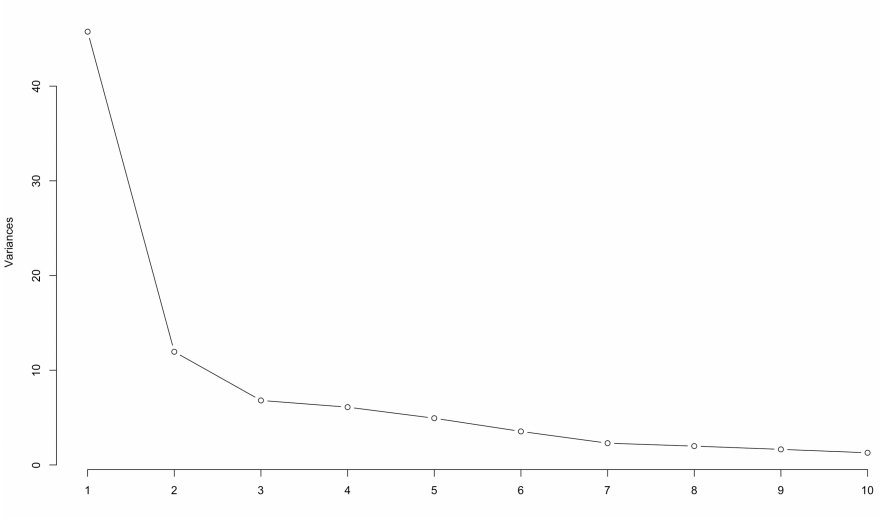


Figure 12: Screeplot for the contribution to overall data variance of the principal components up to PC10

The next step is to examine the relationship between the original dimensions in the data (i.e., the 96 product sectors) and the reduced dimensionality represented by PC1 and PC2. Figure 13 displays the loadings for PC1 and PC2, that is the weights with which each of the 96 regulated product types contributes towards PC1 and PC2. That allows to make sense of the core dimensions summarized by PC1 and PC2. The length of arrows indicates the weight with which each original column contributes to the PCs, while their direction whether they contribute positively or negatively.

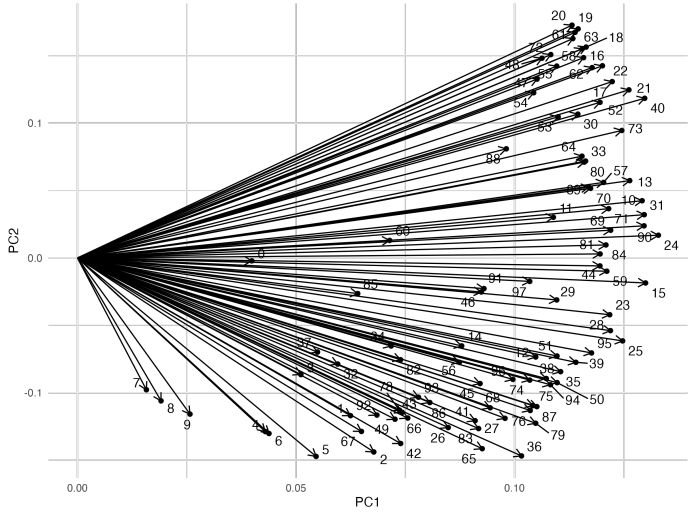


Figure 13: Loadings plot for each product code, 01 to 96 HS chapters.

The plot reveals that country regulations of all 96 product types concur positively to PC1, although to varying degrees. On this basis, the PC1 is first and foremost interpretable as a measure of regulatory intensity, that is of how many NTMs a country has in place. Arrows for all sectors contribute positively

towards PC1. Yet, PC1 also seems to profile countries based on what products their regulations target the most, and specifically whether they focus more on manufactured industrial goods as opposed to agricultural products.

Sectors pointing towards the right-hand section of the plane are those that approximate PC1 most closely. Machinery (84) is almost perfectly parallel to PC1, closely approached by metals (81). The sector with the overall highest loading towards PC1 (around 0.17) is tobacco products (24), closely followed by medical devices (90). Wood, wood products (44) as well as textile fabrics (59) are also part of this cluster. Agricultural and animal products are consistently the sectors with the lowest weights towards PC1. Specifically, the arrows for agricultural goods like vegetables (07), fruit (08), animal products (05), as well as coffee and tea (09), describe an angle of 45 to 60 degrees with the PC1 axis. Their weights range between 0.02 and 0.05. In sum, countries that regulate more heavily manufactured products will have higher PC1 scores than countries regulating agricultural products more intensively.

Moving to the second principal component (PC2), certain sectors contribute positively to the PC2, while others negatively so. This division is not straightforward to interpret. For example, preparations of vegetables (20) is the category with the highest positive weight towards PC2 (above 0.15), closely followed by cereals (19), while agricultural products, as well as trees and plants, score between -0.10 and -0.15 (06, 07, 08, 09). While this distribution is theoretically puzzling, it only accounts for around 10% of total data variance, unlike PC1 which alone accounts for about 50%. Therefore, PC1 emerges as the most reliable measure of a country's regulatory ideal point, as it both captures most of the variance in the data and is also more easily interpretable than the other principal components.

C Regulatory Distance: Validation

As a last step, this section presents a few checks to validate the robustness of my regulatory distance as a measure, including by confronting it to other existing measures. First, I compute a different specification of the simple distance between PC1 scores, relying on the Mahalanobis distance between the PC1 scores of two countries in a dyad. Unlike simple absolute distance, Mahalanobis distance takes into account the overall distribution of PC1 scores in the data, considering its mean and variance. On this basis, it weighs raw distance scores based on the spread of the data. The Mahalanobis distance specification seems particularly appropriate considering how the distribution of PC1 raw distances in the dataset is heavily right-skewed, with most country dyads clustering around or just above the zero, thus displaying minor distance levels.[11] Mahalanobis distance weights raw distance values to ensure that even small variations in distance are accounted for as

[11] See Figure 4.

meaningful variations, while down-weighing unusually large distance changes.[12] In so doing, Mahalanobis distance qualifies as a conservative measure of distance especially in the context of a longitudinal analysis like the one presented in this study, as it avoids that only temporary or isolated drops in PC1 distances pull statistical results. I compute the Mahalanobis distance for each country dyad in my dataset in each year. Figure 14 (below) plots the association between regulatory distance and regulatory Mahalanobis distance.

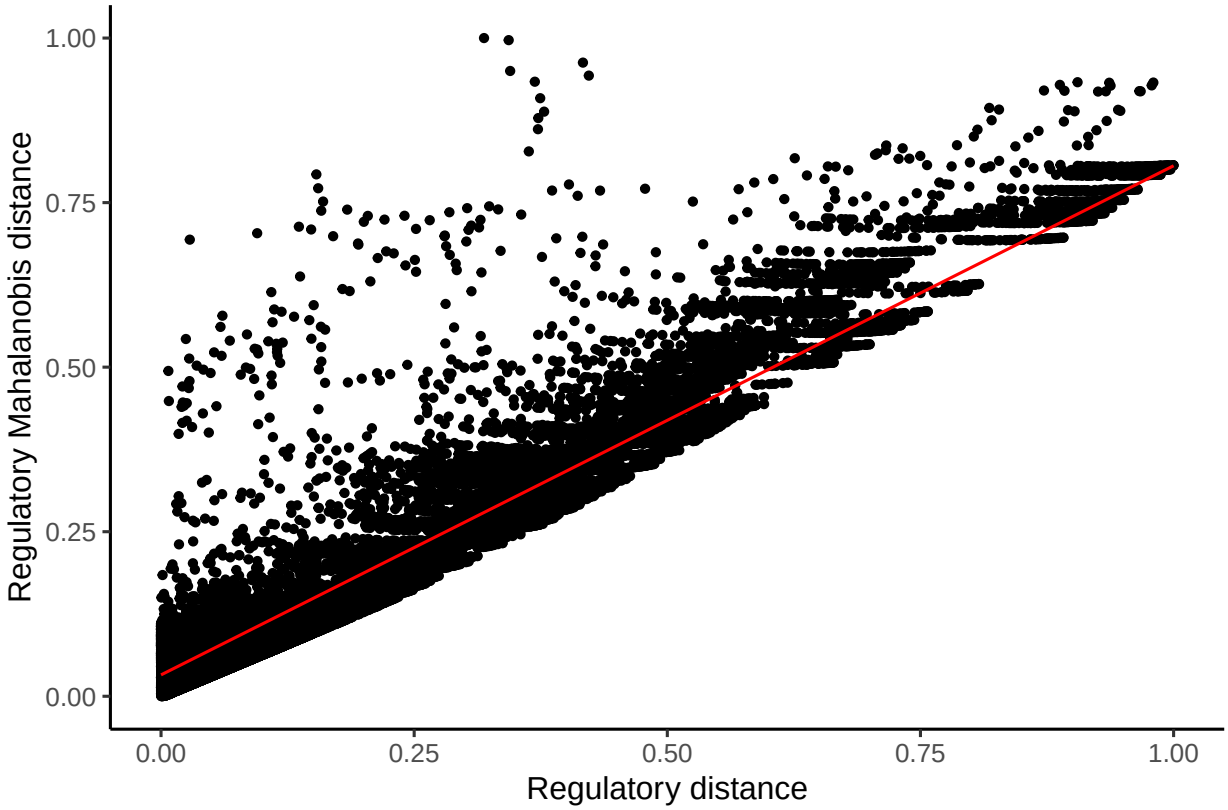


Figure 14: Visual association between Regulatory Distance and Regulatory Mahalanobis Distance.

The above scatterplot suggests a positive and linear association between the two measures of regulatory distance. This strong positive association is also supported when performing a Pearson correlation test, which yields a coefficient of 0.94, statistically significant at the 1% level.

A second way of testing the validity of my regulatory distance measure is confronting it with other existing measures having tried to compute the regulatory proximity of country dyads. The only existing study having done so is by Garcés and Vogt (2024) and presents cross-sectional regulatory distance measures for country dyads in the year 2016. In so doing, it provides snapshot of regulatory proximity in that year, although it does not compute such measure longitudinally as proposed here. Still, it seems useful to confront

[12]The Mahalanobis distance is defined as $d_M(\vec{x}, Q) = \sqrt{(\vec{x}-\vec{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\vec{x}-\vec{\mu})}$. It weights raw differences by the data's variance-covariance structure.

the distance measures for the year 2016.

The authors measure regulatory distance along 3 dimensions: regulatory intensity, coverage, and structure. For the purpose of comparing my measure, I consider their distance measures of intensity, which traces the distance between country dyads based on the number of measures they have in place by product in the year 2016. This seems a suitable measure for comparison considering how the regulatory frequency across products is also the main dimension of interests mapped by PC1 in this study (which does so at the product-family level, HS 2-digit code). Garcés and Vogt (2024) consider the following measure types: TBT, SPS, pre-shipping inspections (PSI), Non-automatic import licensing, quotas, prohibitions, quantity-control and other restrictions, and price-controls, including additional taxes and charges. Their data are also retrieved from the UNCTAD TRAINS Database, which is helpful for comparison purposes.

I confront both my regulatory distance and regulatory Manhattan distance with two measures proposed by Garcés and Vogt for 2016: the i) simple-matching and ii) Jaccard distance in regulatory intensity. The former considers the total number of measure-product combinations imposed by a country of destination in a dyad. Its value decreases with the joint presence or absence of measures for each product, also accounting for textual similarities between measures. In contrast, Jaccard distance only decreases if two countries actually have a product-level measure in common. Running Pearson correlation tests for the year 2016, my measure of simple regulatory distance is 0.12 correlated with the simple-matching intensity distance (significant at the 1% level) and 0.27 (significant at the 1% level) with the Jaccard intensity distance from Garcés and Vogt. My Mahalanobis distance measure is 0.18 correlated with the simple-matching intensity distance (significant at the 1% level) and 0.20 (significant at the 1% level) with the Jaccard intensity distance.

The above correlation scores generally indicate a moderate but consistently positive and significant relationship between the distance measures presented here and those proposed by Garcés and Vogt. These results seem promising if one considers that my dataset disaggregates regulatory measures only by product family (HS 2-digit code), rather than individual products as Garcés and Vogt do. This is an important difference considering how, in the original data matrix, disaggregating intensity at the product level would lower measure frequency under each dimension (i.e., the product) and the likelihood for two countries' similarity under each product. In contrast, disaggregating intensity by product family reduces the number of dimensions of interest, hence also increasing the likelihood that two countries align more closely in their regulatory intensity. That said, the measure I propose also consider a more conservative number of regulations compared to Garcés and Vogt, crucially excluding measures like quotas, prohibitions, and price-controls, which are arguably more temporary and are less telling of a country's domestic regulatory profile.

D PC1: Descriptive Statistics (Dyad-level)

Descriptive evidence helps validate my regulatory distance measure both at a dyadic level. At a macro level, global regulatory distance seems to have grown by around 1% since the establishment of the WTO in 1995.[13] This is consistent with the literature arguing that, while tariffs have generally dropped following GATT/WTO agreements, policymakers have increasingly looked at regulatory barriers as tariff substitutes to address political demands for protection (Gulotty, 2020; Kono, 2006; Perlman, 2023). This is also consistent with domestic regulatory trends. Over the the past three decades, large market economies, like the European Union, have established themselves as "regulatory states", expanding their elaborate systems of domestic regulations, including in domains having an impact on trade (Majone, 1994). I begin with Figure 15 (below), which plots the regulatory ideal points (PC1) for member states of the European Free Trade Association (EFTA). [14] While EFTA is not a customs union, its members are highly integrated trade-wise, and often negotiate preferential trade agreements (PTAs) as a single international entity. Furthermore, its members, are close neighbors of the European Union, which is their chief partner market. This makes them equally exposed to the diffusion of EU market standards and regulations. Switzerland, for example is tied with the EU by a PTA and a network of 16 bilateral agreements covering issues ranging from TBT to the free movement of people (Lavenex, 2009).

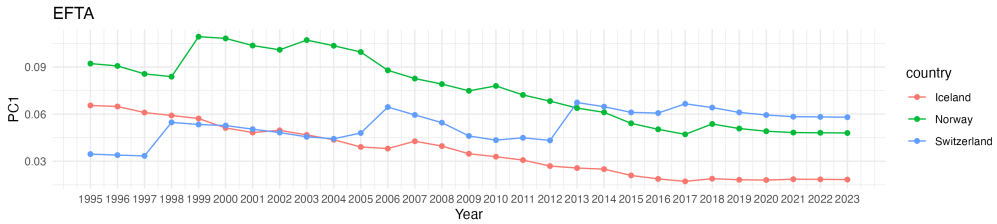


Figure 15: Regulatory ideal points (PC1) for TBT and SPS in EFTA countries (1995-2023)

The trends emerging from the plot are consistent with EFTA countries being closely integrated markets. The maximum regulatory distance recorded between EFTA members amounts to 6% for Norway and Switzerland in 1995, although the two countries have progressively converged, presenting a 1.5% distance by 2023. The higher average PC1 scores for Norway and Switzerland suggest that both countries, being larger markets than Iceland, generally tend to regulate traded goods more intensively. An example of lesser integrated but cooperating markets are the United States and Japan, which have concluded a PTA at the end of 2019. Figure 16 (below) plots the regulatory profiles for these two countries between 1995 and 2023. The highest

[13] See Figure 4.
[14] Liechtestein is excluded since it automatically adopts Swiss NTMs as per the 1923 bilateral Customs Treaty.

regulatory distance between the US and Japan is recorded in 1995 (around 95%). Since then, their distance has progressively reduced, largely as a function of the US PC1 score steadily decreasing since 1995. Their distance has reached a minimum of around 25% in 2019, when the US-Japan PTA was signed, and has been steady until 2023.

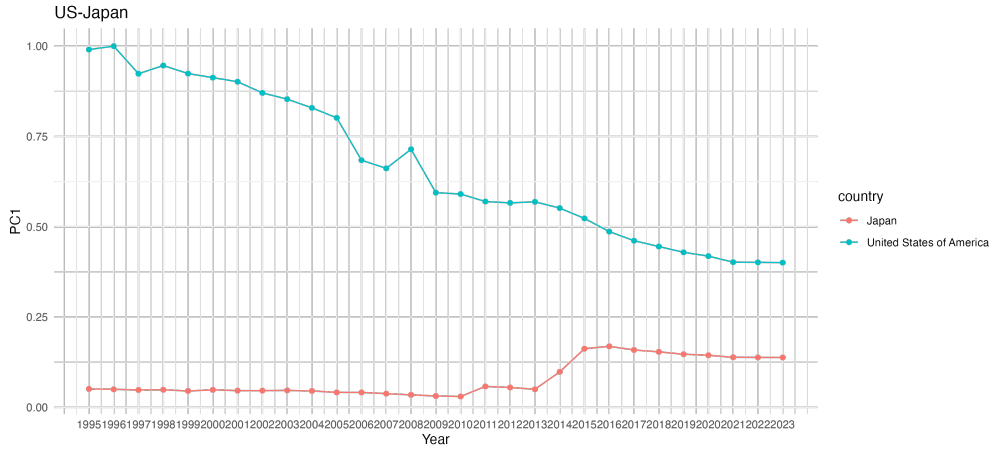


Figure 16: Regulatory ideal points (PC1) for TBT and SPS in the United States and Japan (1995-2023)

E Covariate Balance (DiD)

E.1 Covariate balance and parallel trends for the baseline model (see Figure 7)

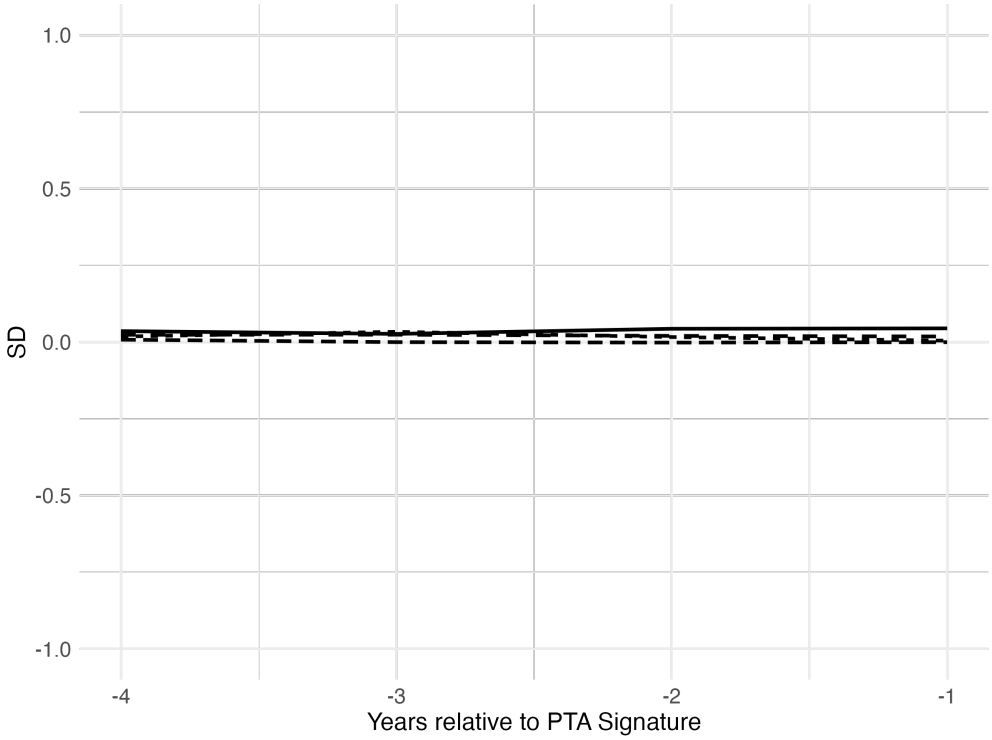


Figure 17: Balance and Parallel Trends using propensity-score weighting refinement in the 4 years preceding PTA signature. The covariate balance is computed across country dyads having concluded a PTA at any point in time between 1995 and 2020.

E.2 Covariate balance and parallel trends, Mahalanobis distance (see Figure 8)

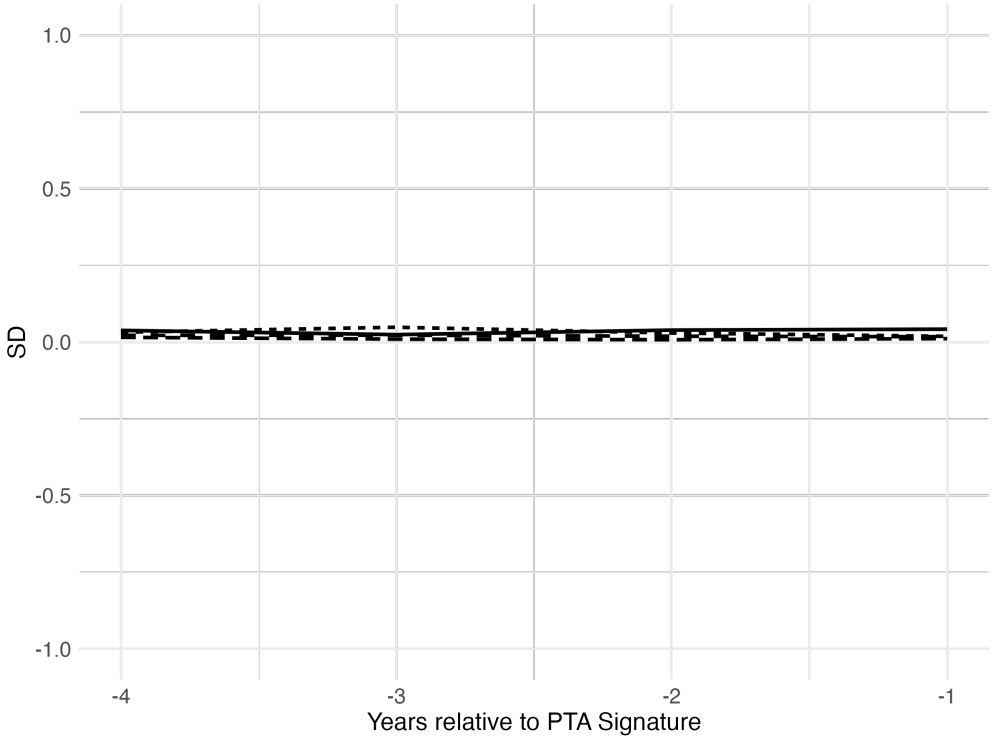


Figure 18: Balance and Parallel Trends using propensity-score weighting refinement in the 4 years preceding PTA signature. The covariate balance is computed across country dyads having concluded a PTA at any point in time between 1995 and 2020.

E.3 Covariate balance and parallel trends for the robustness estimations with PC1 Mahalanobis distance, (see Figure 11)

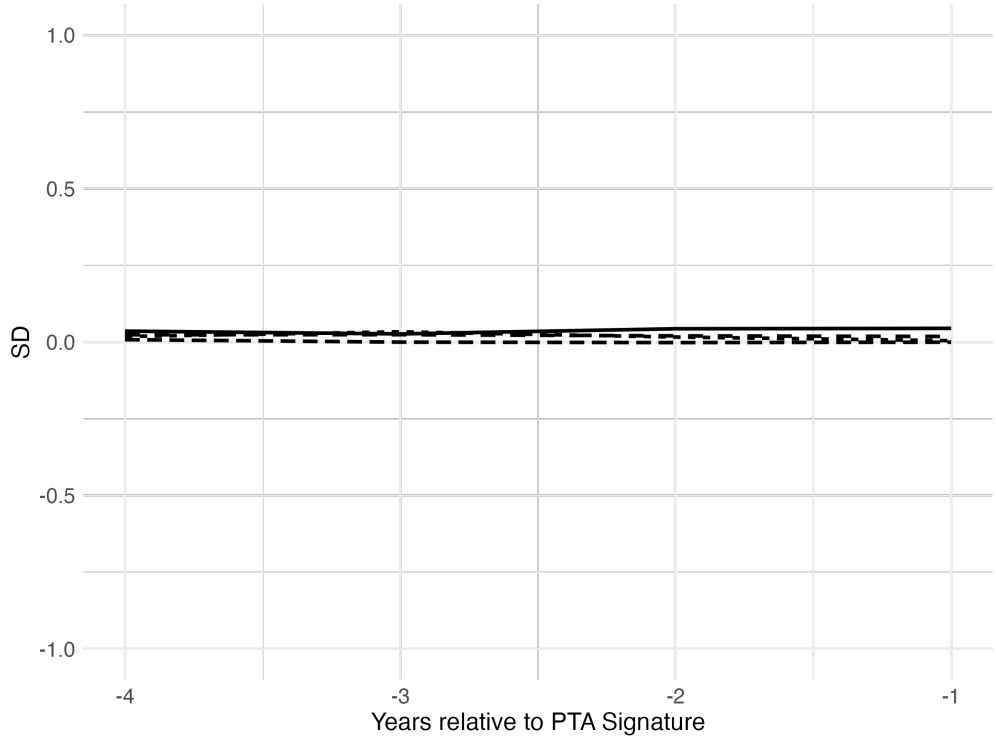
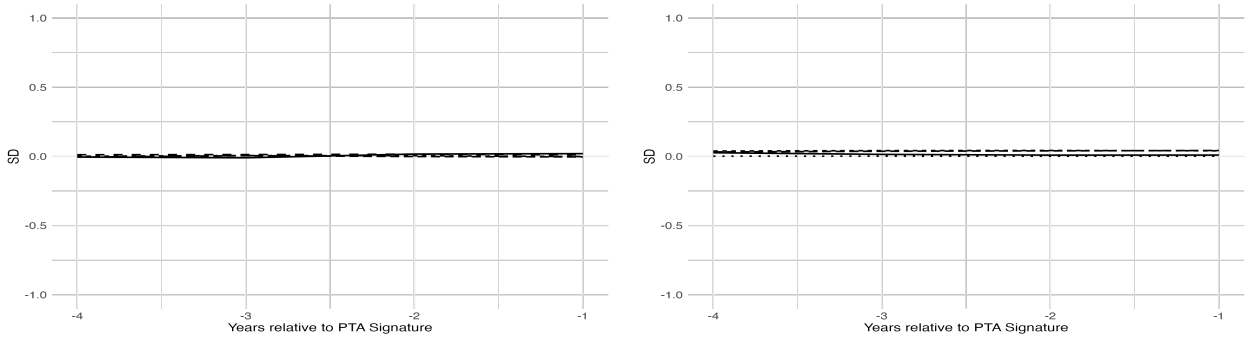


Figure 19: Balance and Parallel Trends using propensity-score weighting refinement in the 4 years preceding PTA signature. The covariate balance is computed across country dyads having concluded a PTA at any point in time between 1995 and 2020.

E.4 Covariate balance and parallel trends for the estimations across the full sample of country dyads, including non-signatories of PTAs (see Figure 9)



(a) DC Institutionalization > Mean **(b)** DC Institutionalization < Mean

Figure 20: Balance and Parallel Trends using propensity-score weighting refinement in the 4 years preceding PTA signature. Covariate balance is computed across all country dyads between 1995 and 2020.