Do Deep Trade Agreements' Provisions *Actually* Increase – or Decrease – Trade and/or FDI?[*]

Jeffrey H. Bergstrand[†] and Jordi Paniagua[‡]

July 18, 2025

Abstract: Over the past 30 years, deep trade agreements (DTAs) have expanded, influencing trade and multinational enterprises (MNEs) through numerous provisions. However, estimating the individual effects of these provisions remains underdeveloped. Using the World Bank's DTA database, this paper applies the Shapley Value approach from cooperative game theory to generate unbiased estimates of the signs of all individual provisions' partial effects. We find that DTAs' provisions can have both positive and negative effects on trade. Our method addresses biases from omitted variables, over-aggregation, and multicollinearity. Beyond trade, we introduce a new dataset on MNEs, examining provisions' effects on bilateral FDI, costs, employment, revenues, and assets. We provide strong evidence that trade-boosting provisions tend to reduce FDI, and vice versa, indicating a substitution relationship with respect to DTAs' provisions. Finally, we present computable general equilibrium welfare estimates for various policy counterfactuals.

Keywords: International trade, foreign direct investment, multinational enterprises, foreign affiliate sales, deep trade agreements

JEL Classification Numbers: F1, F21

[*] *Acknowledgements*: This research was conducted as part of the Project PID2021-122133NB-I00 financed by MCIN/AEI/ 10.13039/501100011033/FEDER, EU, and CIAICO/2021/006 financed by Generalitat Valenciana, and the Kellogg Institute (University of Notre Dame).

[†] *Affiliation:* Department of Finance (Mendoza College of Business), Department of Economics (College of Arts and Letters), and Helen Kellogg Institute for International Studies (Keough School of Global Affairs), University of Notre Dame, and CESifo Munich. *Address:* Mendoza College of Business, University of Notre Dame, Notre Dame, IN 46556 USA. *E-mail:* bergstrand.1@nd.edu.

[‡] *Affiliation:* Department of Economics, University of Valencia and Keough School of Global Affairs and Helen Kellogg Institute for International Studies, University of Notre Dame. *Address:* Department of Applied Economics II, University of Valencia, Avenue Tarongers s/n, 46022 Valencia, Spain. Corresponding author *E-mail:* jordi.paniagua@uv.es

1 Introduction and Motivation

More than fifteen years ago Baier and Bergstrand (2007) asked the question: "Do Free Trade Agreements *Actually* Increase Members' International Trade?" The purpose of that study was to motivate and implement a panel-data econometric specification for providing unbiased and precise partial effects on bilateral aggregate international trade flows of the formation of free trade agreements (FTAs), which could then be used for numerical calculation of economic welfare effects of the implementation (or elimination) of an FTA, cf., Head and Mayer (2014), United States International Trade Commission (2016) and United States International Trade Commission (2021). This methodology became widely adopted and was even employed for numerous policy analyses of FTA formations and dissolutions. However, following Baier and Bergstrand (2007), a plethora of studies have each estimated a single partial "average treatment effect" to summarize the effect of the presence or absence of an economic integration (or preferential trade) agreement.[1]

However, the past 30 years have witnessed the proliferation of "deep trade agreements," or DTAs. Preferential trade agreements (PTAs), such as free trade agreements, historically focused on reductions of tariff rates across a broad swath of product categories. The recent use of the word "deep" refers to the evolution of PTAs over the last 30 years. As Pascal Lamy, former Director General of the World Trade Organization, wrote in the foreword to the World Bank's *Handbook of Deep Trade Agreements* (2020, p. x):

> *"Deep trade agreements... are fundamentally different than the previous generation of PTAs. They aim not only to create market access between members but also to establish broader economic integration rights in goods, services, and factor markets. Deep agreements support these rights by regulating the behavior of importing and exporting governments."*

Lamy suggests that DTAs are different from earlier PTAs in two basic dimensions. First, DTAs deepen international liberalization of goods trade beyond tariff cuts by reducing both border and "behind-the-border" barriers as well as *broaden the liberalization* of international exchange of services, capital, people and ideas. Second, DTAs also increasingly provide and enforce provisions that regulate the behaviors of exporters and importers (in some cases, to meet non-trade goals of governments); such provisions are explicit *obligations* in DTAs and can potentially reduce trade and foreign direct investment flows. The term DTAs is synonymous with the term economic integration agreements (EIAs), which has been used elsewhere in the literature, cf., Baier et al. (2014), Bergstrand et al. (2015), and Baier et al. (2018).

[1] Furthermore, this methodology has also been used for generating unbiased and precise estimates of the well known "trade elasticity," that is, the elasticity of bilateral trade flows with respect to a tariff-rate change, and consequently the numerical calculations of the "gains from trade," cf., Arkolakis et al. (2012).

In this paper, we offer four potential contributions. First, because of the enormous number of provisions in modern DTAs, the measurement of such provisions' partial effects has evolved over time. Early analyses used the *number* of provisions to measure the depth of the agreement and the consequent effects of DTAs on trade flows. More recent analyses have employed various machine-learning techniques to categorize sets of provisions. However, simple counts and machine-learning approaches mask positive-versus-negative effects of individual provisions within their aggregations, leading to misleading quantitative impacts and consequent inferences. In this paper, we draw instead upon the Shapley Value approach from cooperative game theory to guide us toward estimating the effects of sets of provisions on trade versus foreign direct investment (FDI). This approach allows us to provide estimates of the *positive-versus-negative* effects of provisions on trade and on FDI, an aspect that has been largely overlooked in the literature, while also addressing methodologically the challenges posed by omitted-variables bias, over-aggregation bias, and multi-collinearity. While not able to generate an unbiased estimate of each individual provision's partial quantitative effect, Shapley Values do generate unbiased estimates of the *signs* of individual provisions' effects, which we then separate into small sets of types of provisions. Among numerous results, we provide *ex post* evidence that "liberalizations" can have positive *or negative* effects on trade and "obligations" can have negative *or positive* effects.

Second, while this is not the first study to examine the trade-flow impacts of DTA provisions, this is the first study (to the authors' knowledge) aimed at estimating partial effects of sets of DTAs' provisions on members' foreign direct investment of multinational enterprises (MNEs) *as well as on other relevant measures of MNEs' activities* including foreign affiliates' costs, revenues, employment, and assets (alongside members' trade flows). To do so, we introduce a new data set, the Multinational Revenue, Employment and Investment Database, or MREID. While trade economists now have sufficient-statistics approaches to summarize economic welfare effects of provisions, MNEs themselves (along with policymakers) would find informative estimated effects of such provisions on costs, revenues, employment, and assets.

Third, our study is novel by examining the effects on FDI of provisions that positively affect trade, and the effects on FDI of provisions that negatively affect trade. We find evidence that sets of provisions that positively (negatively) affect trade also negatively (positively) affect FDI. This is consistent with evidence in Bergstrand and Egger (2007) and Mistura and Roulet (2019) that free trade agreements positively affect trade, but negatively affect FDI. Furthermore, while such gravity models have been used to estimate the effects on bilateral FDI stocks of various bilateral costs of FDI, we provide novel bilateral gravity-equation methodological enhancements to show empirically that most investment provisions have *multilateral* effects and that DTA provisions have (third-country) *spillover* effects.

Finally, the range of DTA provisions provides a rich source of potential policy counterfactuals. Consequently, we also provide some representative numerical computations of the general equilibrium effects of timely counterfactuals, using state-of-the-art sufficient-statistics approaches.

The remainder of the paper is as follows. In Section 2, we provide an initial brief overview of data on DTAs (section 4 will provide a more detailed and novel overview focusing on the MREID database). We review the extant literature on estimating DTAs' partial effects on trade. While there are a few studies addressing DTAs' impacts on foreign direct investment (FDI) and foreign affiliate sales (FAS) revenue, none additionally examine DTAs provisions' impacts on other measures of MNEs' activities such as costs, employment, and assets.

In Section 3, we briefly summarize the theoretical framework of trade, FDI, and FAS undergirding our structural gravity equation econometric approach.

In Section 4, we discuss in detail the three data sets undergirding our empirical analysis. First, the DTA database of the World Bank provides binary indicators of nearly 1,000 DTA provisions. For our purposes, the World Bank has provided a subset of "substantive" provisions. Such provisions – at the very heart of the two fundamental directions of DTAs relative to traditional PTAs – provide a unique opportunity to explore empirically the positive – and potentially negative – effects of liberalizations and obligations on trade and MNEs' activities. Second, we introduce a novel data set on multinational firms' "activities." The new data set is the Multinational Revenue, Employment, and Investment Database (MREID), summarized in Ahmad et al. (2025). This database is the most comprehensive one to date with information on MNEs' FDI-related activities with cross-border affiliates across the pairings of 185 countries, 25 industries, and 12 years (with annual updates) at the bilateral *transaction* level. Compiled using data from Orbis, this data set has detailed observations of MNEs' activities that is usually constrained to a single source, such as the U.S. Bureau of Economics Analysis database on U.S. MNEs' transactions; our database includes 185 countries' MNEs' bilateral transactions. Moreover, it contrasts with the OECD's or IMF's databases on bilateral FDI levels which is only available for a large number of country pairs at the bilateral *aggregate* level. Among numerous advantages, our data set has several distinguishing features such as the inclusion of measures of *domestic* (or intra-national) investment (necessary for appropriate structural gravity estimation). Third, we review the trade data from the established International Trade and Production Database for Estimation v2 (ITPD-E).

In Section 5, we discuss our empirical strategy. The underlying framework uses the well-established structural gravity equation econometric specification developed in Baier and Bergstrand (2007) and Baier et al. (2014) for estimating such treatment effects. However, the multitude of provisions warrants examining methodological considerations of omitted variables bias (OVB) and over aggregation bias (OAB), alongside the imprecision of coefficient estimates (and potential amplification of omitted variables bias) from multi-collinearity (MC). As an alternative to recent machine-learning techniques, we introduce the Shapley Value approach from cooperative game theory to correct for the biases associated with OVB and OAB, while still addressing concerns over MC. We use a simulated world of six countries with bilateral trade, bilateral agreements, and various "liberalization" and "obligation" provisions to demonstrate how our procedure reduces bias in estimates of positive or negative effects of narrow sets of provisions.

In Section 6, we provide the econometric specifications and main estimation results, using the Poisson Pseudo-Maximum-Likelihood estimator recommended in Santos Silva and Tenreyro (2006) and adapted for three-way high-dimensional fixed effects in Weidner and Zylkin (2021). Among numerous novel results, we now briefly note some key findings. First, we find evidence that DTA provisions positively affect trade (FDI), with one additional (randomly selected) provision in the World Bank's data set (with nearly 1,000 provisions) increasing *aggregate* trade (FDI) by 0.06 (0.01) of 1 percent. Second, the marginal effect of a "substantive" provision on trade (FDI) is four (three) times that of a non-substantive provision. Third, DTAs with the mean number of substantive provisions (30) increase trade (FDI) by 11 (2) percent. Fourth, substantive provisions have quadratic effects on trade and FDI, but the nature of the quadratic relationship differs between them. Fifth, when all individual provisions that have positive (negative) effects on FDI are grouped, these sets of provisions positively (negatively) affect FDI *but* negatively (positively) affect trade. Furthermore, when all individual provisions that have positive (negative) effects on trade are grouped, these sets of provisions positively (negatively) affect trade *but* negatively (positively) affect FDI. These results are all statistically significant. Sixth, we find unbiased estimates of the signs of individual provisions' effects on trade and FDI, noting that (the World Bank's) "liberalization provisions" can have positive or negative effects and their "obligation provisions" can have negative or positive effects. Seventh, owing to the breadth of data in our MREID database, we identify channels through which such effects work; for example, we find statistically significant evidence that sets of provisions that positively (negatively) affect FDI are decreasing (increasing) marginal *costs per employee* at affiliates.

In Section 7, we provide an extensive robustness analysis of the main empirical findings. We show that our results are qualitatively similar whether we use our "Grouped Shapley" approach or our "Individual Shapley" approach, Poisson Pseudo-Maximum Likelihood or Ordinary Least Squares estimation, and aggregate bilateral data on foreign affiliates or averages per affiliate. This section also distinguishes between estimates of the "bilateral" versus "multilateral" partial effects of provisions and provides evidence of "spillover" effects.

In Section 8, we provide several general equilibrium comparative static exercises to provide numerical welfare impacts of changes in certain provisions. Among numerous results, we show that inclusion of provisions with only "negative trade" effects lowers world output by 0.77%. By contrast, inclusion of provisions with only "positive trade" effects raises world output by 2.7%.

Section 9 provides conclusions.

2 Overview of DTAs and Relevant Literature

Section 2.1 provides a brief overview of the Deep Trade Agreements (DTA) Database of the World Bank; a more detailed (and novel) discussion is provided in section 4. Section 2.2 discusses previous estimates of the effects of different aggregations of the DTA Database provisions on trade. The last subsection discusses the estimates of FTas and DTAs on FDI.

2.1 Measurement of Deep Trade Agreements

The earliest study to evaluate systematically the depth of modern EIAs was Horn et al. (2010). Horn et al. (2010) evaluated the "content" of 14 European Union (EU) and 14 United States (US) EIAs. First, they introduced the distinction between agreements that were WTO+ versus WTO-X. WTO+ refers to the content of an EIA that is consistent with the current WTO mandates. For example, an EIA that lowers tariff rates below Most Favored Nation (MFN) rates is consistent with WTO+. By contrast, WTO-X refers to a provision in an EIA that goes outside (at that time) the current WTO mandates, such as an *obligation* to conform to a common labor standard. Second, they examined whether the obligations they observed were legally enforceable or not. Third, they found systematic evidence that US agreements contained more legally enforceable WTO-X provisions than EU agreements, and US agreements have put more emphasis on legally enforceable provisions in *regulatory areas*. A useful survey of these issues is provided in Limao (2016).

The first paper released by the World Bank summarizing characteristics of the World Bank's DTA Database was Hofmann et al. (2017); this working paper introduced the DTA Database to the world. The DTA Database provides a 0-1 indicator variable response (No/Yes) to 972 questions regarding "provisions" across 52 broad policy categories among 279 economic integration agreements (EIAs) bilaterally for pairings of 189 countries annually from 1958 through 2017. The 52 policy categories – which overlap with "WTO+" and "WTO-X" provisions – are provided in Table 1 of Hofmann et al. (2017).[2] The database

[2] Hofmann et al. (2017) typically use the term "preferential trade agreement" (PTA) to refer to agreements. However, given the breadth of liberalization among flows of goods, services, capital, people, and ideas, we use in this paper the term Economic Integration Agreement, or EIA.

provides the most horizontally extensive in breadth – and vertically extensive in depth – database recognizing the presence or absence of provisions related to the 52 policy categories.

In an analysis of "stylized facts" pertaining to the DTA Database, Mattoo et al. (2020) note several important dimensions. First, measuring the "depth" of 279 EIAs is no trivial task. At one end, two of the three early approaches measured the depth of an EIA using simple "count" measures. For instance, Hofmann et al. (2017) defined *TotalDepth* of an EIA as simply the number of provisions found. These authors also defined a measure *CoreDepth* as the number of "core" provisions, where a core provision is defined as one in any of the first seven of the 19 areas listed in the paragraph above. The third early measure, *PCADepth*, uses a weighted sum of provisions where the weights are determined by a principal components analysis. At the other end, some studies have employed "machine-learning" techniques to measure an EIA's depth, c.f., Breinlich et al. (2021) and Baier and Regmi (2023).

Second, as Mattoo et al. (2020) note, although the "number of provisions" is a very useful approximation for the depth of an EIA, some provisions are more "substantive" or "essential" than others. The authors note that the distinction between substantive and non-substantive provisions is important. Moreover, an important issue overlooked is that some measures have positive effects on trade while others have negative effects. We show later that the marginal effect of one positive-trade-impact substantive provision has *eight times* the impact of one randomly chosen substantive provision.

2.2 Previous Estimates of Effects of DTAs on Trade

The measurement of the quantitative (partial) impacts of EIAs on bilateral trade, FDI, and FAS began *systematically* only 15 years ago, and focused predominantly on average treatment effects of free trade agreements (FTAs) on bilateral trade flows. Prior to 2007, naive gravity-equation econometric specifications (predominantly using cross-sections) suffered unknowingly from severe endogeneity biases, leading to widely varying – and often economically implausible – average treatment effects. For international trade and free trade agreements (FTAs), Baier and Bergstrand (2007) is widely acknowledged as the standard econometric approach for estimating empirically partial effects of trade agreements on trade flows. Using ordinary least squares (OLS), panel data, and three-way fixed effects, Baier and Bergstrand (2007) provided econometrically consistent and economically plausible partial (average treatment) effects of FTAs on bilateral trade flows – with a typical effect (coefficient in a log specification) of 60 percent (0.46). Building on that study and its methodology, Baier et al. (2014) examined specifically the partial effects of various degrees of "depth" of EIAs on trade flows (i.e., using separate dummies for one-way PTAs, two-way PTAs, FTAs, customs unions, common markets, and economic unions) and also examined the differential effects of the types of EIAs on the extensive and intensive margins of trade.

Baier et al. (2018) extended this work to examine the heterogeneous effects of various types of EIAs in terms of depth and interacted with (proxies for) factors that influence variable versus fixed trade costs. Furthermore, building upon the Knowledge-Capital model in Markusen (2002), Bergstrand and Egger (2007) provided the first theoretical general equilibrium Knowledge-and-Physical-Capital model that provided a rationale for estimating structural gravity equations of bilateral trade, foreign direct investment, *and* of MNEs' foreign affiliate sales. One of the quantitative results from a numerical version of that model was that regional *trade* agreements should increase (decrease) bilateral trade (foreign affiliate sales) of members; their empirical specifications confirmed those predictions.

More recent work has employed this methodology in various manners to expand the literature on "depth" of EIAs, and such measures' effects on bilateral trade flows. Mattoo et al. (2017) employs the DTA Database to determine whether the depth of agreements – measured by three alternative provision *count* variables – impacts trade flows. The authors used the methodology of Baier and Bergstrand (2007) in a Poisson Pseudo-Maximum Likelihood (PPML) estimator and found that all three measures of the *numbers* of provisions influenced trade flows; the earliest measure of depth using the DTA Database has been various "counts" of provisions. For instance, Mattoo et al. (2017) used (alternatively) a count of all provisions, a count of weakly legally enforceable provisions, and a count of legally enforceable provisions. A distinguishing element of Mattoo et al. (2017) is that the authors also accounted for non-discriminatory aspects of numerous deep trade provisions, implying less potential trade diversion; in fact, the positive spillover effects on non-members implied negative trade diversion effects. Kohl et al. (2016) similarly essentially uses counts of various provisions under WTO+ and WTO-X and estimates the impacts of these count variables on bilateral trade flows; however, to avoid multicollinearity, these authors include count variables for various policy areas *separately*. Mulabdic et al. (2017) also employed a count measure of DTA provisions to analyze UK-EU trade relationships.[3]

Fontagne et al. (2022) estimated the partial effects of DTAs on trade. However, instead of the count approach for measuring depth, Fontagne et al. (2022) introduced a "classification algorithm" that grouped – what they argued were – similar provisions into "clusters." In their case, they grouped 910 provisions into three clusters. They employed a well-established classification method known as the "k-mean++" clustering algorithm.[4] The major goal of creating groups or clusters of provisions is to avoid the severe multi-collinearity associated with including – in their case – 910 dummy variables for each of the DTA database's provisions.

[3]Dhingra et al. (2018) focused on a single measure of services, investment, and competition policy provisions provided by the DESTA database. The authors found significant effects for their clustered variables, but could not address the effects of individual provisions.

[4]As stated in Fontagne et al. (2022), "This iterative algorithm partitions the data into a number of pre-determined clusters based on a dissimilarity matrix measuring the Euclidean distance between agreements across the 18 policy areas covered in each of the 278 treaties under analysis" (p. 6).

However, clustering into only three groups – analogous to the principal components approach in Mattoo et al. (2017) – makes it difficult to understand *which* provisions country-pairs might want to utilize in forming an EIA. In reality, MNEs' lobbyists are generally interested in very specific provisions – or, at least, policy areas – for their respective industries. Such lobbyists argue their cases in public testimony, which ultimately filters into the formation of the EIAs.More recently, Herman (2024) estimates the non-linear effects of the number of DTA provisions on trade, concluding that DTAs that are "too deep" hinder trade.

Breinlich et al. (2021) also estimated the partial effects of DTAs on trade using classification algorithms. In that paper, the authors used a LASSO (Least Absolute Shrinkage and Selection Operator) method for provision selection. The key idea in their approach is to "shrink" the number of provisions – by statistically eliminating the ones with zero trade effects – so that the number of provisions represented is well below the number of EIAs. In their study, the number of provisions selected is too large for the model to have meaningful interpretations. Hence, to address the limitations of the LASSO technique, they introduce two further methods: an "iceberg" LASSO and also a "bootstrap" LASSO. In the end, the iceberg LASSO suggests a set of 42 provisions that impact trade and the bootstrap LASSO identifies 30-74 provisions. While their results usefully narrow the number of impactful provisions to six (one for anti-dumping, one for competition policy, two for technical barriers to trade, and two for trade facilitation), the authors note the possibility that these provisions were selected – not because they were the "real cause" of the estimated trade flow increase, but – because "they tend to appear in EIAs together with other provisions that are the real cause." For instance, the study's (Table 5 Post-Lasso) results suggest that the inclusion of anti-dumping Provision 14 (a requirement to establish material injury to domestic producers) should increase *aggregate* bilateral trade by 42 percent. Such impacts seem highly implausible quantitatively for individual provisions' impacts on aggregate bilateral trade since most recent PPML estimates find partial effects for entire PTAs at approximately 25% (exp(0.22)). Nevertheless, the authors are upfront in stating, "These methods address difficulties arising from the high degree of correlation between (the hundreds of) individual provisions... Though to be clear, they do not completely answer the question of '*Which provisions matter for trade?*' "(p. 2; italics added).[5]

[5] Baier and Regmi (2023) also used machine-learning techniques to estimate the effects of DTAs on trade flows. In that paper, the authors use a simple form of machine learning to identify clusters of provisions that are grouped together. The authors observe the likelihood of each provision in each cluster. Their approach informs them of which provisions – when bundled together – are associated with higher trade flows.

2.3 Previous Numerical Estimates of effects of either FTAs or DTAs on FDI

By contrast with the evolving – and very recent– literature estimating the effects of DTAs on bilateral trade flows using alternative techniques from dummies for individual provisions to complex classification algorithms, there are very few analogous efforts to examine the effects of DTAs on MNEs' bilateral FDI and FAS and *none* that examine partial effects on other variables of relevance to MNEs' behaviors (which we will examine). As noted above, Bergstrand and Egger (2007) found empirical evidence that FTAs reduced FDI between country-pairs. Mistura and Roulet (2019) also find that FTAs reduce FDI flows. By contrast, Buthe and Milner (2014) showed that FTAs increased FDI. Paniagua et al. (2015) reconcile these mixed findings of FTAs on FDI using quantile regressions. They show that FTAs (and BITs) have a positive and significant effect on FDI above the median and negative and insignificant effects in the lower quantiles.

Regarding the deepness of DTAs, Grounder et al. (2019) found that the depth of provisions in DTAs did not significantly affect pairs of countries' FDI. By contrast, Kox and Rojas-Romagosa (2019) showed that the deepest PTA (with an index of seven in the DESTA database) positively affected FDI stocks. Laget et al. (2021) find a positive relationship between greenfield FDI and the number of disciplines (or groups of provisions) in DTAs. Larch and Yotov (2025) obtain a small positive and statistically significant estimate on the number of provisions included in DTA on FDI stocks. They also find negative effects of the number of provisions in certain policy areas.[6]

3 Theoretical Considerations

Our econometric methodology and empirical work are based upon recent theoretical "New Quantitative" models of international trade and multinational enterprises' (MNEs') foreign affiliate sales. In the Online Appendix, we review in detail that such theoretical models are based upon either the homogeneous firms framework or the heterogeneous firms framework. We review theoretical foundations with homogeneous firms established in Bergstrand and Egger (2007), Bergstrand and Egger (2010), Bergstrand and Egger (2013), Bergstrand and Egger (2014), and Anderson et al. (2019). We also review the complementary theoretical foundations with heterogeneous firms in Ramondo and Rodriguez-Clare (2013) and Arkolakis et al. (2018). Both sets of foundations motivate similar structural gravity equations for

[6] Building on the literature on the international organization of production (Antràs and Helpman, 2008), Osnago et al. (2019) find that provisions enhancing the contractibility of components foster FDI, whereas those improving contractibility of headquarters' services reduce it. This highlights the importance of provision content in determining their impact on FDI.

estimation of sets of provisions' effects on trade, FDI, and FAS. Moreover, we review the literature on the international organization of production, noting that provisions in DTAs that improve the contractibility of components may have different effects on FDI than those that improve the contractibility of headquarters' services, implying some provisions may have positive (negative) effects on FDI. Analogously, some provisions – such as trade liberalizations (obligations) – may increase or decrease trade.

Closely related to our study is the theory developed by Anderson et al. (2019) who provided a theoretical framework for the determinants of bilateral trade *and* FDI (assuming homogeneous productivities across firms). Anderson et al. (2019) added two key features to the Anderson and van Wincoop (2003) framework to generate a structural Armington model with physical capital accumulation, "technology capital" accumulation, and bilateral FDI (though no foreign affiliates or FAS).[7] First, they introduced a two-tiered Cobb-Douglas production function for the single national output with the lower tier a Cobb-Douglas function of (internationally immobile) labor and physical capital and the upper tier a function of the lower tier and the (internationally mobile) "global technology stock." Second, the global technology stock for any country j was a Cobb-Douglas function of its own "domestic" technology stock and bilateral FDI from all non-j countries, where each flow from i to j (FDI_{ij}) is subject to a factor representing bilateral FDI openness (ω_{ij}).

Ignoring here the transition dynamics of the model, they generated a system of steady-state equations (for the "upper-level equilibrium") that included the well-known structural gravity model of trade with multilateral resistance terms and four more equations determining the price of each country's national output, each country's level of output, each country's level of expenditures, and each country's level of private physical capital, *as well as* a structural gravity equation for every pair of countries' bilateral FDI stock (FDI_{ij}); the last equation was multiplicative in country i's output, country j's expenditures, and the variable determining openness to i's FDI into j (ω_{ij}). Larch and Yotov (2025) apply Anderson et al.'s 2019 framework to analyze general equilibrium effects on FDI of DTAs. Their findings suggest that DTAs have contributed to a large but very asymmetric increase in inward versus outward FDI.

Note that the models in Bergstrand and Egger (2007) and Anderson et al. (2019) both imply multiplicative (in levels) gravity equations for bilateral trade *and* bilateral FDI, with the Bergstrand-Egger model also implying a multiplicative gravity equation for FAS.

These theoretical contributions, alongside studies cited, provide motivation for gravity equations of bilateral trade and various measures of bilateral FDI estimated in sections 6 and 7.

[7] Preferences were defined by the standard Armington preferences with CES utility.

4 Data

The data for our analysis come from three data sets: the DTA database of the World Bank, a new Multinational Revenue, Employment, and Investment Database (MREID) at the U.S. International Trade Commission (see Ahmad et al. 2025), and the International Trade and Production Database for Estimation v2 (ITPD-E) at the U.S. International Trade Commission (see Borchert et al. 2022b). We discuss each database in turn.

4.1 World Bank's Deep Trade Agreements Database

Section 2 provided a brief overview of the DTA database. In this section, we report in graphical detail some of the characteristics of this data that will be useful to know for our econometric analysis later. First, given the large scope of this paper – descriptive analysis, econometric analysis, and numerical counterfactuals – we chose to limit our analysis to 164 provisions denoted by the World Bank as "Substantive Provisions."[8] Substantive (also called essential) provisions were decomposed at the World Bank into "liberalization/integration" provisions and "obligations/conditions" provisions, cf. (Mattoo et al., 2020, p. 11). However, we will show in section 6 applying our Shapley Value methodology that "liberalization/integration" provisions actually can have (*ex post*) positive or negative partial effects (on trade and FDI) and "obligations/conditions" provisions also can have (*ex post*) positive or negative partial effects, an important consideration toward estimating average treatment effects of narrow groups of provisions.

In the Online Appendix, we provide an extensive description of the DTA database with world maps of concentrations of all substantive provisions, time series of provisions' adoptions, and natures of provisions (bilateral versus multilateral). [9]

4.2 The Multinational Revenue, Employment, and Investment Database (MREID)

In a recent survey of the effects of International Investment Agreements (IIAs), Egger et al. (2023) note the limitations of current measurements of FDI. The existing databases from UNCTAD, IMF, and OECD are known to have heterogeneous reporting standards, share little information on sector flows, and do not adjust for tax-haven behavior of MNEs (see Casella et al. (2023) for a recent survey on FDI datasets). Guvenen et al. (2022) show that

[8] In Table O.2 of the Overview in Mattoo et al. (2020), of the 937 total individual provisions, 243 are listed as Substantive. However, we consider only 13 of the 16 policy groupings, omitting Labor Standards, Movement of Capital (which has no substantive provisions), and Environmental Standards; this yields 164 substantive provisions.

[9] We provide additional visualizations of DTAs in this link: `https://public.tableau.com/app/profile/nabeel.saad/viz/shared/4G3KPMZ5T` .

accounting-engineering practices such as profit-shifting motives are common among U.S. affiliates and impact the measurement of aggregate FDI variables. Another measure of FDI commonly used is a panel of bilateral FDI "projects" compiled by FDIMarkets (under the Financial Times). However, this bilateral data is based upon *projected* (or anticipated) events and not actual transactions; for further details, see Egger et al. (2023). Another additional limitation of current FDI databases is that they provide data on only *foreign* investments. However, the standard state-of-the-art methodology for econometric analysis of treatment effects in the trade and FDI literature uses a structural gravity framework, cf., Bergstrand et al. (2015), which requires comparable data on *domestic* investment as well.

In collaboration with the U.S. International Trade Commission's (USITC's) Office of Economics, the authors created a new database of MNEs' "activities," cf., Ahmad et al. (2025). The Multinational Revenue, Employment, and Investment Database (MREID) provides comprehensive and consistent information on international *and domestic* bilateral revenues, costs, employment, numbers of affiliates, and investment variables of MNEs for the pairings of 185 countries, across 25 industries, and (initially) 12 years.[10] Covering a wide range of agriculture, mining, energy, manufacturing and services industries, MREID provides a novel and comprehensive panel of sectoral-level bilateral foreign direct investment (and domestic investment) and foreign affiliate sales activities; we use the term "foreign direct investment" (FDI) broadly for now, but also narrow the definition later. Furthermore, MREID distinguishes greenfield investment from merger and acquisition (M&A) investment. MREID currently covers the period 2010-2021, with planned annual updates using reported administrative data from ORBIS.

We use a search strategy on the Orbis database to overcome several limitations of previous databases. Orbis is Bureau van Dijk's flagship-company database with data from more than 425 million companies worldwide. It focuses on private company information and presents companies' variables in a comparable formats; information is sourced from over 170 different providers but is standardized into comparable cross-country information.[11]

Orbis is a popular resource among economists. Kalemli-Ozcan et al. (2015) (revised 2023) were the first to describe the standard benchmark-search strategy to construct nationally representative firm-level data from the Orbis global database. Using this search strategy, Gopinath et al. (2017) studied capital stock (fixed assets), output (sales), and employees. These authors show that Orbis data coverage is comparable to Spanish administrative data. Osnago et al. (2019) used Orbis to construct an FDI dataset for several European countries

[10] Our database is unique in providing domestic investment data *compatible with* the FDI data.

[11] The MREID database we construct consists of publicly owned and privately owned corporate firms with assets or sales larger than USD 1 million; hence, most are publicly owned. It excludes banks and state-owned enterprises and banks. FDI requires ownership of 50.01 percent or larger. International generally accepted accounting standards are used.

and were able to distinguish vertical and horizontal FDI. Garcia-Bernardo et al. (2017) used Orbis data to unravel offshore financial centers.

The key variable to foreign identity ownership in Orbis is the variable "Global Ultimate Owner" (GUO), which allows us to bypass some of the offshore issues that plague official FDI statistics of major sources, such as that of the IMF. This variable allows us to track firms that invest in foreign countries. One of the limitations of the Orbis web interface is that the GUO variable is only available for the "current" day. To overcome this limitation, we followed the approach in Kalemli-Özcan et al. (2023) and used the M&A module in Orbis to track these changes over time. With this procedure, we obtained accurate FDI data without accessing historical days; the consequent limitation is a 10-year rolling period.[12]

The dimensions of our database are as follows. MREID (initially) spans 12 years, 2010-2021. It contains financial data of 362,845 parent companies (GUOs) with 1,132,707 domestic and foreign affiliates. Of those, 351,66 are foreign affiliates of 70,661 parent companies, and the rest are domestic affiliates. Raw data from 25 sectors are combined and, after undergoing data cleaning, we have approximately 27,000 raw observations per year at the country-sector (two-digit) level.

MREID has data on FDI for 186 countries; hence, there are potentially 34,410 (=186x185) bilateral FDI "measures" of activity. However, FDI data are characterized by a large number of zeros; hence, the raw MREID database is an unbalanced panel.[13]

Table 1 provides summary statistics associated with 4,817 country-pairs with a least one foreign affiliate investment. The mean number of active foreign affiliates across country-pairs (and averaged over 2010-2021) in the sample is 90.

Table 2 reports summary statistics on (time-averaged) revenues, employees, and total and fixed assets by ownership (i.e., domestic vs. foreign). Domestic affiliate statistics include all affiliates of parent companies from the same country. As discussed earlier, only 139 countries in the sample report domestic affiliates. Countries have 5,869 active domestic affiliates, on average. Foreign affiliate statistics include all affiliates of parent companies from different countries; hence, statistics in Table 2 (Panels A and B) are at the country level. As expected, aggregate values are higher for domestic than foreign affiliates.

Table 3 reports summary statistics on (time-averaged) revenue, number of employees, and total and fixed assets *per affiliate* and *by ownership* (i.e., domestic vs. foreign). Note that the average foreign affiliate tends to be larger in (per affiliate) revenues, number of

[12]This procedure allows construction of a comparable companion dataset recording M&A data. Whenever an affiliate enters the database within the observation period of 2010-2021, it is flagged as a greenfield investment; hence we also have a second companion database on greenfield investment. We limited our search to affiliates with more than USD 1 million in sales or in total assets in at least one year in the sample. We also implemented criteria to detect exits from the market. Ahmad et al. (2025) provides extensive details on the search strategy.

[13]For our estimates, we construct a square dataset including zeros.

employees, and assets than the domestic one. Moreover, the largest foreign affiliates (max) are larger than the domestic ones in (per affiliate) revenues, number of employees, and fixed assets.

4.3 The International Trade and Production Database for Estimation v2 (ITPD-E)

The third dataset used is the International Trade and Production Database for Estimation (ITPD-E) at the U.S. International Trade Commission. The second version of this dataset, as

Table 1: Summary Statistics for Foreign Affiliates at the Country-Pair Level

	Panel A: Totals			Panel B: Avg. per Affiliate		
	Mean	Max	SD	Mean	Max	SD
No. of For. Affiliates	90	25,299	536			
Revenue	3,940	609,312	20,362	57	5,772	236
Employees	7,029	1,735,375	43,965	191	156,239	2,619
Total Assets	14,480	6,309,828	132,300	221	56,616	1,432
Fixed Assets	5,198	1,615,221	48,817	60	15,276	507
Revenue/Employee	48,251	65,794,332	1,282,092			
N					4,817	

Notes: N is number of country-pairs with foreign affiliates; SD is standard deviation.
In both panels, revenue and total and fixed assets are in million USD.
In Panel A, revenue per employee is in thousands of USD.
In both panels, employees denotes the actual number.

Table 2: Summary Statistics at the Host Country by Ownership (Totals)

	Panel A: Domestic			Panel B: Foreign		
	Mean	Max	SD	Mean	Max	SD
No. of Affiliates	5,869	164,199	19,246	1,984	54,430	5,687
Revenue	136,628	3,570,717	471,000	86,441	1,666,594	238,122
Employees	246,864	4,783,207	764,243	152,329	3,968,938	482,269
Total Assets	763,302	28,438,464	3,351,904	316,189	12,108,262	1,174,622
Fixed Assets	132,133	5,199,483	540,606	113,942	4,000,906	473,700
Revenue/Employee	1,029	21,801	2,667	3,583	227,384	21,773
N		139			175	

Notes: Revenue and assets are in millions of USD. Revenue/employee are in thousands USD.
Foreign statistics are at the host country level.
N denotes number of countries in the sample and SD the standard deviation.

described by Borchert et al. (2022b), encompasses data from 265 countries, 170 industries, and spans 33 years (1986-2019). The underlying data for ITDP-E come from several sources that provide clear documentation, contain data that were not estimated by statistical procedures, and that are regularly updated. These sources include FAOSTAT, COMTRADE, and WTO for international trade and FAOSTAT, MINSTAT, INDSTAT, and the UN National Account Statistics for domestic production.

An essential characteristic of the ITPD-E is that it contains international and domestic trade flows, analogous to MREID containing international and domestic FDI measures. The combined use of MREID and ITDP-E for FDI and trade, respectively, allows for comparable and theory-consistent gravity estimations. Borchert et al. (2021, 2022a) run standard gravity regression using the ITPD-E dataset[14] and show consistent and well-behaved estimates of the usual gravity variables.

The combined use of both datasets, however, imposes several limitations in terms of country and time coverage. For our analysis, we used a subset of the MREID and ITPD-E that assures the presence of domestic and international data for the maximum number of countries and years. This renders a dataset with 138 countries that spans 2010 through 2019. Table 4 reports the summary statistics for the left-hand-side variables used in our estimates.[15]

Table 5 reports the summary statistics for the right-hand-side variables in our dataset. Panel I shows the statistics of all country pairs, where we can appreciate that 19.5% of country pairs in our sample have a DTA in force (see top row). Panel II shows the statistics

Table 3: Summary Statistics at the Host Country by Ownership (per Affiliate)

	Panel A: Domestic			Panel B: Foreign		
	Mean	Max	SD	Mean	Max	SD
Revenue	76	970	171	93	1,224	188
Employees	250	3,829	624	282	5,095	697
Total Assets	424	11,394	1,224	431	5,505	749
Fixed Assets	51	1,490	160	94	3,915	428
N		137			172	

Notes: Revenue and assets are in millions of USD.
Foreign statistics are at the host country level.
N denotes number of countries in the sample.

[14] They use the v1 dataset, which is limited to data until 2016.

[15] Some of the variables, like costs and assets, are not publicly available in the first release of the MREID dataset. However, they come from the same source and were compiled with the same procedures described in Section 4.2.

of country pairs that have a DTA in force with at least one substantive provision. We observe that DTAs in our sample contain nearly 30 substantive provisions, on average.

5 Methodology: Econometric Problems and Solutions

In principle, we would like to be able to provide an unbiased and consistent estimate of the partial effect of each *individual* provision on bilateral trade and FDI.[16] However, as we address below, the literature on estimating effects of provisions on trade flows has been plagued by concerns over omitted variables bias (OVB), multicollinearity (MC), and over aggregation bias (OAB). To make headway, in section 5.1 we set up an illustrative – or "true" – model of bilateral trade flows as a function of individual provisions with (initially) homogeneous effects. In this section, we demonstrate quantitatively using our simulated world the biases caused by omitted variables and multicollinearity. In section 5.2, we allow our stylized model

Table 4: Descriptive Statistics of Dependent Variables

	Mean	SD	Min	Max	Units
Trade	4424.3	182238.3	0	25366110.1	Million USD
FDI (number of affiliates)	45.47	1420.8	0	163913	Number
Employee Costs	1021919.0	10530676.4	0.00480	656991234.5	Thousand USD
Cost per Employee	371.3	8535.3	0.00000357	814378.3	Thousand USD
Employees	2790.4	68677.1	0	6157044.1	Number
Tangible Assets	471045.8	14339034.6	0	1.96954e+09	Thousand USD
Intangible Assets	133920.6	4680586.6	0	811586712.4	Thousand USD
Revenues	1580736.2	41027474.3	0	4.48858e+09	Thousand USD
Observations	190440				

Notes: Statistics are at the country-pair level.

Statistics are for 138 origin and destination countries during 2010-2019.

Not all firms report employee costs and the observations are limited to 18784 observations.

Table 5: Descriptive Statistics of Provisions

	Mean	SD	Min	Max	Mean	SD	Min	Max
	Panel I				Panel II			
DTA (dummy)	0.195	0.396	0	1				
All provisions	46.60	96.61	0	595	165.1	116.6	16	595
Non-substantive provisions	38.73	80.94	0	480	137.1	98.91	13	480
Substantive provisions	7.871	16.00	0	120	27.95	18.64	2	120
Observations	190440							

Notes: Panel I reports statistics using the full dataset.

Panel II reports statistics of those country pairs with a DTA in force with at least one provision.

[16] Such an estimate could then be used for an estimate of its general equilibrium effect.

to have heterogeneous provisions to examine the influence of over-aggregation bias (OAB). In section 5.3, we show how a Shapley Value approach can provide a bias correction for estimating the average positive and average negative effects of (narrowly defined) groups of provisions; Shapley Values can only provide unbiased estimates of the *signs* of individual provisions' partial effects. In the Online Supplement, we assess quantitatively our method with simulations on a cross-section of the dataset. Our methodology has a success rate of over 90 percent in identifying correctly the assignment of individual provisions into narrowly defined groups for which average treatment effects can be estimated, even allowing *ex ante* heterogeneity among such individual provisions' effects.

5.1 A Stylized Model of DTAs with Homogeneous Provisions

To illustrate the problems of OVB, MC, and OAB in our context, we consider a hypothetical world of only six countries; it will be useful for tractability, however, to assign the names Spain (ESP), Germany (DEU), United States (USA), Canada (CAN), Brazil (BRZ), and Argentina (ARG). For simplicity, we assume all six countries are of equal size economically and we ignore non-policy (or "natural") trade and investment costs. We assume that the United States and Canada belong to a shallower trade and investment agreement, which we label DTA-1. We assume that Spain and Germany belong to a DTA, which we label DTA-2. We assume Brazil and Argentina belong to a trade agreement, which we label DTA-3.

For the exercise at hand, we assume DTA-1 is composed of two provisions. Assume D_1 is a provision that can either increase or decrease (foreign direct) investment flows between USA and CAN and D_2 is a provision that liberalizes their international trade. We will consider two cases in what follows. In Case 1, we consider the case where the investment provision (D_1) *doubles* trade and the export liberalization (D_2) also doubles trade between trading partners. Later in Case 2, we allow investment provision D_1 to *halve* trade.[17]

We assume the DTA-2 is a deep "trade" agreement that includes four provisions. D_1 is an investment (INV) provision in the EU between ESP and DEU; we assume it is of the same magnitude as in DTA-1 and has the same quantitative impact as between CAN and USA. D_2 enhances exports (EXP), liberalizing trade between ESP and DEU; we assume it is of the same magnitude as in DTA-1 and has the same quantitative impact. D_3 is a competition-policy (CPP) provision that we assume (initially) doubles ESP-DEU bilateral trade, but has no effect on their FDI. D_4 is a technical-barriers-to-trade (TBT) provision that (initially) doubles their bilateral trade, but has no effect on their FDI. We assume the agreement between Brazil and Argentina (DTA-3) includes only trade provisions D_2 and

[17] Case 2 (heterogenous provisions) seems the more intuitive case. However, Case 1 (homogenous provisions) is used initially to provide a quantitative contrast to Case 2.

D_4 with the same magnitudes of effects. The rationale for introducing CPP and TBT is to allow later for heterogeneous trade impacts of provisions.

Table 6 contains descriptive statistics and the (partial) correlation-coefficients matrix for the provision dummies relevant to our Case 1 (homogenous provisions) with Spain, Germany, Canada, the United States, Argentina and Brazil, which will prove useful shortly. Note that the variation in the mean values of each variable reflects the variance of inclusion of the provision among the agreements.

Table 6: Descriptive Statistics and Correlation Matrix

			(1)		
	Mean	SD	D_1 INV	D_2 EXP	D_3 CPP
D_1 INV	0.222	0.422	1		
D_2 EXP	0.333	0.478	0.756***	1	
D_3 CPP	0.222	0.422	0.357*	0.756***	1
D_4 TBT	0.111	0.319	0.661***	0.500**	0.661***

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, N=36.

For brevity, we focus here only on trade-flow impacts of provisions; the FDI impacts follow analogously. Suppose X_{ijt} denotes the level of trade from exporter i to importer j in year t and ϵ_{ijt} is an error term with mean 1. We assume the true model of trade is:

$$X_{ijt} \quad = \quad \left(\prod_{p=1}^{P} e^{\beta_p D_{p,ijt}} \right) \epsilon_{ijt}. \tag{1}$$

where $D_{p,ijt}$ is a dummy variable that takes the value 1 in period t if the country-pair ij has provision p, and 0 otherwise. The "true" effect of provisions is unknown to the econometrician, who faces a challenge estimating the effects of provisions and DTAs.

5.1.1 Omitted Variables Bias and Multicollinearity Bias

Initially (for Case 1 - homogenous), we assume each of the four provisions *doubles trade* (including INV); hence, the expected coefficient estimates in the absence of omitted variables bias (OVB) and multicollinearity (MC) is approximately 0.7. As well known, OVB can be substantial. For instance, consider the specification including only a single provision (D_1):

$$\ln X_{ij} = \tilde{\beta}_1 D_1 + \tilde{\epsilon}_{ij}. \tag{2}$$

The expected value of $\tilde{\beta}_1$ is given by:

$$E\left[\hat{\tilde{\beta}}_1\right] = \beta_1 + \sum_{p=2}^{4} \beta_p \rho_{1n} \frac{\sigma_{D_p}}{\sigma_{D_1}} \tag{3}$$

where ρ is the correlation coefficient between a pair of dummy variables and σ_{Dp} denotes the standard deviation of D_p, which are taken from Table 6. Given OVB and the correlation matrix in Table 6, the expected coefficient estimate for $\tilde{\beta}_1$ is positive rather than negative.

Multicollinearity (MC) is a well-known problem when multiple regressors are linearly correlated. Since a regression coefficient is intended to measure a change in one variable with all other variables constant, if two or more regressors are linearly correlated, the interpretation of the coefficient estimates becomes very difficult. We will demonstrate shortly, using our stylized model, that near-perfect multicollinearity can lead to biased coefficient estimates.

We use our stylized model to evaluate OVB and MC with simulated trade data between countries. Initially, trade data is i.i.d. $X \sim N(100, 1)$. Table 7 reports the estimates for individual DTAs provisions, where we illustrate both biases. Columns 1–4 of Table 7 show the effects of OVB by estimating each provision separately.[18] As predicted, the coefficients of all provisions are biased (with the true coefficient 0.693), highlighting the relevance of the OVB in this case. Column 5 of Table 7 highlights the effects of MC. Investment and CP provisions are dropped due to perfect multicollinearity. D_2 (EXP) picks up the pro-trade effects of D_1 (INV) and D_2 (EXP). The TBT provision picks up the total effect of CP and TBT.

Many earlier analyses have included individual provisions both in studies of trade effects and FDI effects. However, as shown, there is little to be learned about the individual effect of each provision if either OVB or MC are present.

5.2 Over Aggregation Bias

As just shown, little is to be learned by including individual provisions' dummies alone (OVB) or all the provisions simultaneously (MC). Consequently, the literature has introduced using the *number* (or count) of provisions to avoid these issues. We show in this section that including a simple count of provisions provides an unbiased estimate of the marginal effect of a provision *when all provisions have identical effects*. However, a simple count variable can provide a very biased estimate if provisions' effects are heterogeneous.

[18] For these simulations, we use ordinary least squares (OLS) as we assume no zero trade flows and no heteroskedasticity. Consequently, the results using OLS or Poisson Pseudo Maximum Likelihood (PPML) are identical. In empirical work later in sections 6 and 7, we use PPML.

In the context of our stylized model (and assuming positive flows here, for simplicity), we can rewrite equation (1) as:

$$\ln X_{ij} = \beta_1 D_{1,ij} + \beta_2 D_{2,ij} + ... + \beta_P D_{P,ij} + \ln \epsilon_{ij} \tag{4}$$

where, for convenience, we omit the time (t) subscript. In the case of homogeneous effects, $\beta_1 = \beta_2 = ... = \beta_P = \beta_{SP}$, we can write:

$$\ln X_{ij} = \beta_{SP} \sum_{p=1}^{P} D_{p,ij} + \ln \epsilon_{ij} \tag{5}$$

or

$$\ln X_{ij} = \beta_{SP} SPROV_{ij} + \ln \epsilon_{ij} \tag{6}$$

where $SPROV_{ij} = \sum_{p=1}^{P} D_{p,ij}$ and β_{SP} is the *marginal* effect of one more provision on trade.

Table 7: Simulated World: Omitted Variables Bias and Multicollinearity Issues

	(1)	(2)	(3)	(4)	(5)
D_1 INV	1.879***				0.000
	(0.22)				(.)
D_2 EXP		1.847***			1.384***
		(0.14)			(0.00)
D_3 CPP			1.881***		0.003
			(0.22)		(0.00)
D_4 TBT				2.424***	1.384***
				(0.31)	(0.00)
Observations	36	36	36	36	36
R^2	0.68	0.84	0.68	0.64	1.00
$\exp D_1$	$\times$ 6.55				$\times$ 1.00
$\exp D_2$		$\times$ 6.34			$\times$ 3.99
$\exp D_3$			$\times$ 6.56		$\times$ 1.00
$\exp D_4$				$\times$ 11.29	$\times$ 3.99

Notes: Standard errors are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2.1 Case 1: Homogeneous Provisions

In the context of Case 1 of our stylized model where every one of the four provisions *doubles* trade, we can readily show that a simple count of provisions provides unbiased estimates. In this case, the expected value of β_{SP} is 0.693.

Table 8 confirms this along with additional insights. First, $\hat{\beta}_{SP} = 0.693$ in the bottom row of the top panel of Table 8, as expected. Second, we also gain insights from coefficient estimates for dummy variables for DTA-1 (D_{DTA-1}), the DTA-2 (D_{DTA-2}), and DTA-3 (D_{DTA-3}). In our model's setup in Case 1, INV and EXP each double trade. One can see from the bottom panel that the two provisions together increase trade fourfold ($\exp D_{DTA-1} = 3.99$). For the European Union, there are four provisions, each doubling trade, so the total effect is sixteenfold ($2^4 = 16$), and the coefficient of 2.771 is expected ($\exp 2.77 = 15.97$). Analogously, Mercosur has two trade provisions (EXP and TBT), each of which doubles trade; hence, ($\exp 1.387 = 4.00$). Finally, we note the result in the top row of Table 8. DTA is a dummy variable for any pair ij that has a provision. Not surprisingly, the coefficient estimate for DTA is in between the values for D_{DTA-1}, D_{DTA-2}, and D_{DTA-3}.

Table 8: Effects of DTAs and of Numbers of *Homogeneous* (Positive) Provisions

	(1)	(2)	(3)
DTA Dummy	1.847***		
	(0.14)		
D_{DTA-1}		1.384***	
		(0.00)	
D_{DTA-2}		2.771***	
		(0.00)	
D_{DTA-3}		1.387***	
		(0.00)	
SPROV			0.693***
			(0.00)
Observations	36	36	36
R^2	0.84	1.00	1.00
$\exp D_{DTA}$	×6.34		
$\exp D_{DTA-1}$		× 3.99	
$\exp D_{DTA-2}$		× 15.97	
$\exp D_{DTA-3}$		× 4.00	
$\exp SPROV$			×2.00

Notes: Standard errors are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2.2 Case 2: Heterogenous Provisions

The bias associated with "over-aggregation" arises if we allow a provision to have a "heterogeneous" effect from the others. Consider now Case 2, where the only distinction vis-a-vis Case 1 is that the foreign direct investment provision (INV) *halves trade*. This provision is present in both NAFTA and the European Union.

Table 9 provides the results of estimating our model under alternative specifications. Notably, in contrast to Table 8, the marginal effect of one more provision increases trade by *half* of the corresponding value in Table 8 (0.346 vs. 0.693). Why? There are two reasons. Beyond just the change from INV doubling trade in Case 1 to halving trade in Case 2, what matters also is the presence or absence of the INV provision *across pairs* of countries. Clearly, the diminution of $\hat{\beta}_{SP}$ from Case 1 to Case 2 is related to the presence of a provision with a *different effect* from the others. Here, INV now has a negative effect, but we could easily have simply changed INV's trade effect to a value *other than* doubling.

Not surprisingly, in Table 9, the coefficient estimate for the DTA dummy variable decreases from 1.847 in Case 1 to 0.923 in Case 2. Furthermore, since NAFTA includes INV that halves trade and includes EXP that doubles trade, NAFTA does not have a net effect on trade (exp D_{NA} =1). Analogous interpretations follow for the European Union and Mercosur.

The resulting important question becomes: Is there a method to determine the expected estimate of β_{SP} in the presence of heterogeneous provisions?

5.3 Shapley Value Methodology

The theoretical foundation behind the novel technique we propose is inspired by the Shapley Value (SV) imputation from the cooperative game-theory literature, cf. Lipovetsky and Conklin (2001). As discussed earlier, multicollinearity (MC) is a serious problem when multiple regressors are highly correlated. Theoretically important variables may have insignificant coefficients; the results in Table 7 confirm this. Lipovetsky and Conklin (2001) specifically address formally the use of the SV approach from game theory to address regression analysis in the presence of extremely high MC among regressors. While the application in their paper deals specifically with variables that all have positive *ex ante* effects on the regressand, the same methodology extends to our case of interest of positive and negative effects of provisions.

The essence of the SV imputation from game theory is that it interprets the utility of a particular "player" over *all possible combinations of players.* In the context of a regression model, the SV imputation assigns a value for each "predictor" (or variable) calculated over all possible combinations of predictors in a regression. Intuitively, by comparing across all possible combinations, it captures the possibility of the competitive influences of any subset of predictors.[19] The canonical SV technique, following Darlington (1968) and Harris (1975), considers the measure of the relative importance of predictors and consists of evaluating the usefulness of each regressor via the *increment* to the multiple determination of a model (R^2).

However, two problems surface in our particular regression application with regard to the strict use of R^2 as a measure for SV imputation. First, the workhorse estimation technique for structural gravity equations is Poisson pseudo maximum likelihood (PPML); however, PPML does not report an R^2 but rather a Pseudo-R^2.[20] Second, we are not interested in the predictive power of individual provisions but rather their *marginal* (and potentially causal) effect.[21] To address these concerns, we use a two-step approach.

Table 9: Effects of DTAs and Numbers of *Heterogeneous* (Negative and Positive) Provisions

	(1)	(2)	(3)
DTA Dummy	0.923***		
	(0.14)		
D_{DTA-1}		-0.002	
		(0.00)	
D_{DTA-2}		1.384***	
		(0.00)	
D_{DTA-3}		1.387***	
		(0.00)	
SPROV			0.346***
			(0.04)
Observations	36	36	36
R^2	0.84	1.00	0.68
exp D_{DTA}	×2.52		
exp D_{DTA-1}		×1.00	
exp D_{DTA-2}		× 3.99	
exp D_{DTA-3}		× 4.00	
exp $SPROV$			×1.41

Notes: Standard errors are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[19] The leave-one-out cross-validation technique in Bayesian models has a similar intuition (Vehtari et al., 2017).

[20] The Pseudo-R^2 is the correlation between the predicted values from the PPML estimation and the observed values.

[21] Causal inference is a major drawback for most alternative maximum likelihood techniques.

5.3.1 First Stage

After initially estimating $\hat{\beta}_{SP}$ using equation (6), we estimate a regression with two variables. One variable is the dummy for a particular individual provision, which we denote provision n. The second variable is $SPROV$ (defined earlier) excluding (or *minus*) provision n. Formally (and, for simplicity, expressed in logs and omitting the time t subscript), the regression is:

$$\ln X_{ij} = \beta_n \cdot D_{n,ij} + \beta_{SP-n} \cdot (SPROV - D_n)_{ij} + \ln \epsilon_{ij}. \tag{7}$$

However, the coefficient estimate for D_n is *not* the Shapley "Value"; it is referred to as a Shapley "coefficient estimate." We will demonstrate shortly – using a simulation – two problems with using the actual estimate of β_n, as Lipovetsky and Conklin (2001) discuss. The Shapley coefficient estimates may be inflated. Moreover, the coefficient estimates may not even capture correctly the *sign* of the effect of a provision on trade (or FDI); Shapley *values*, however, will provide unbiased estimates of the sign of the effect.

Consequently, to first retrieve unbiased estimates of the (positive vs. negative) *signs* of the effect of a provision, we turn, as in Lipovetsky and Conklin (2001), to the Shapley "values." In our context, the Shapley Value (SV) of individual provision n is:

$$SV(D_n) \equiv \underbrace{\hat{\beta}_{SP}\overline{SPROV}}_{\hat{\beta}_{DTA}} - \underbrace{\hat{\beta}_{SP-n} \cdot (\overline{SPROV - D_n})}_{\hat{\beta}_{DTA-n}} \tag{8}$$

where $\overline{SPROV} = \sum_{ij=1}^{N} \left(\sum_{p=1}^{P} D_{p,ij} \right) /N$ where N is the total number of country-pairs ij with provisions and $\overline{SPROV - D_n} = \sum_{ij=1}^{N} \left(\sum_{p=1}^{P} D_{p,ij} - D_n \right) /N$. Recall that $\hat{\beta}_{SP}$ was determined earlier using equation (6) and $\hat{\beta}_{SP-n}$ can be estimated using equation (7).

We will show later in this section using simulations that the SV approach (first step) provides *unbiased* estimates of the relative importances of multiple RHS determinants of a LHS variable and, in our context, *unbiased* estimates of the positive vs. negative effects of a provision – even with *heterogeneous* individual provision effects.

5.3.2 Second Stage

Although estimates of β_n from equation (7) tend not to be unbiased estimates of individual provisions' quantitative effects on trade or FDI, the Shapley Values determined in equation (8) are useful by providing unbiased estimates of the relative importances of multiple regressors. Consequently, in principle one can "group" various provisions into subsets according to alternative criteria. For instance, one can use this technique to distinguish between "substantive" vs. "non-substantive" provisions without relying upon subjective interpretations. Also, one can use this technique to distinguish provisions that are "liberalizations" vs. "obligations." Moreover, one can group provisions into certain "policy areas."

For the purposes of this study's empirical work later, we will be interested in determining which provisions positively or negatively affect trade and which provisions positively or negatively affect FDI. In fact, we will be able to show that – using the Shapley Value technique – many provisions labeled "liberalizations" by the World Bank have positive (trade or FDI) effects, but many of these have *negative effects*. Analogously, we will show that many provisions labeled "obligations" by the World Bank have negative effects, but many of these have *positive effects*. Furthermore, we can potentially determine the manner in which sets of provisions that positively or negatively affect FDI (or trade) have an effect on *other* bilateral measures of MNE "activities" – such as numbers of employees, costs of employees, costs per employee, tangible assets, intangible assets, and revenues at their affiliates. This will be the focus of our empirical work later.

However, to anticipate some of those findings in our simulations here, our second step will split all substantive provisions $SPROV$ into two groups: all provisions with positive effects on trade ($\ln X_{ij}$) and all provisions with negative effects on trade. Hence, in our second step, we split the $SPROV$ variable into two variables: one is the sum of all provisions with positive effects on trade ($SV(D_n) \geq 0$) and the other is the sum of all provisions with negative effects on trade. [22] Consequently, OAB is addressed by grouping provisions into homogeneous groups (i.e., positive provisions and negative provisions separately). Shortly, we will show that this method provides unbiased estimates of the signs of the individual provisions' average effects on trade – with homogeneous effects or even heterogeneous effects.

The following equation provides a biased-corrected estimate of the effect of individual provisions and the total effect of DTAs by separating positive from negative provisions:

$$\ln X_{ij} = \underbrace{\beta_{SP}^{+} SPROV_{ij}^{+}}_{SV(D_n)>0} + \underbrace{\beta_{SP}^{-} SPROV_{ij}^{-}}_{SV(D_n)<0} + \ln \epsilon_{ij} \tag{9}$$

where $SPROV_{ij}^{+}$ is the sum for each country-pair ij of provisions that positively affect trade (determined by positive SV values) and $SPROV_{ij}^{-}$ is the sum for each ij of provisions that negatively affect trade (determined by negative SV values). Equation (9) shows that the effect of DTAs is composed of a positive element and a negative element. This insight is interesting for policy because, if the ultimate policy goal is to design and enforce DTAs that maximize trade, policymakers should include those provisions that actually increase trade.[23]

[22] We assume that all substantive provisions have a revelant effec on trade and/or FDI.

[23] Since endogeneity bias may arise, we will demonstrate later in the empirical analysis that this concern is mitigated by considering how sets of provisions that positively (negatively) affect trade influence FDI, and how sets of provisions that positively (negatively) affect FDI influence trade.

5.3.3 Simulation Results

Table 10 below reports the results of estimating equation (7) for each of the four individual provisions (D_1-D_4) of our simulated world in columns 1-4 (under Case 2). To interpret the findings, focus first on column (1). In this regression, we are estimating the effect of the INV provision and the sum of the rest on trade. As one can see from column (1), the Shapley regression captures precisely the incremental (negative) contribution of the investment liberalization, with the coefficient estimate of D_1 being approximately -0.7, as expected. Moreover, since all the other three provisions double trade, the coefficient estimate for PROV w/o INV is 0.7. Furthermore, in the bottom panel, the first row indicates the Shapley Value for this regression of -0.464, indicating correctly that this provision reduces trade; hence, INV is labeled a *Negative Trade Provision*. Both the regression coefficient estimates and the Shapley Values are correct.

Consider next column (2). In this case, we are estimating the incremental effect of a trade liberalization, D_2, that should double trade. In this case, the coefficient estimate for D_2 is half the size of 0.7 (which would double trade), which is biased downward. Furthermore, the coefficient estimate is statistically insignificant. However, the SV for column (2) correctly indicates that this provision has a positive effect on trade, and thus is categorized as a *Positive Trade Provision*.

Similar findings hold for CPP (D_3) and TBT (D_4) in columns (3) and (4). As expected, CPP and TBT are *Positive Trade Provisions* based upon their Shapley Values in the top row of the bottom panel of Table 10.

Overall, for columns (1)-(4) the SV of each individual provision captures correctly the sign but is lower (in absolute terms) than the estimates of their individual coefficients.

Finally, consider column (5). In this regression using equation (9), we have combined all *Positive Trade Provisions* into a variable $SPROV^+$ and the one *Negative Trade Provision*, D_1, into $SPROV^-$. Consistent with our theory and discussion above, $SPROV^+$ doubles trade whereas $SPROV^-$ halves trade. Recall, however, that these results are based upon an assumption of *homogeneous* positive effects. We address this limitation next.

In reality, positive trade provisions will have heterogeneous partial effects and negative trade provisions will have heterogeneous partial effects. To illustrate quantitatively the implications for our coefficient estimates and Shapley Values using the previous exercise but now allowing heterogeneous partial effects, we modify the previous exercise so that the three trade provisions (*EXP, CPP, TBT*) have the following (true) partial effects. We assume EXP doubles trade, but CPP increases trade *fourfold* and TBT increases trade *eightfold*. As before, we assume INV halves trade. Table 11 reports the coefficient estimates as well as the Shapley Values in the top row of the bottom panel.

We note three findings of importance. First, in column (1), we note that the coefficient estimate for PROV w/o INV is much larger than before, reflecting that D_3 and D_4 have much larger (assumed) effects. Second, and a standard implication of Shapley Value coefficient estimates, the coefficient estimate for INV is biased upward (in absolute value); it should be -0.693. However, as expected, the Shapley Value for INV in column (1) is negative, and so this implies it is a negative trade provision for subsequent grouping. Third, note that the

Table 10: Coefficient Estimates and Shapley Values (Homogeneous Effects)

	(1)	(2)	(3)	(4)	(5)
		First Stage			Second Stage
D_1 INV	-0.696***				
	(0.00)				
PROV w/o INV	0.693***				
	(0.00)				
D_2 EXP		0.346			
		(0.21)			
PROV w/o EXP		0.346***			
		(0.10)			
D_3 CPP			1.388***		
			(0.00)		
PROV w/o CPP			-0.001*		
			(0.00)		
D_4 TBT				0.346	
				(0.25)	
PROV w/o TBT				0.346***	
				(0.07)	
SPROV$^+$					0.693***
					(0.00)
SPROV$^-$					-0.696***
					(0.00)
$SV(D_n)$	-0.464	0.346	0.925	0.231	
$\exp SV(D_n)$	$\times$ 0.63	$\times$ 1.41	$\times$ 2.52	$\times$ 1.12	
$\exp \beta_n$	$\times$ 0.50	$\times$ 1.41	$\times$ 4.01	$\times$ 1.41	
$\exp \beta^+$					$\times$ 2.00
$\exp \beta^-$					$\times$ 0.50

Notes: Standard errors are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

coefficient estimate for D_2 in column (2) is negative. However, the positive Shapley Value in column (2) correctly indicates that this is a positive trade provision. Using the coefficient of D_2 instead of its SV would incorrectly categorize this provision as negative. The Shapley Values in columns (3) and (4) also correctly indicate that they are positive trade provisions.

Importantly, in the last two rows of the bottom panel of Table 11, we observe that the effect of the positive provisions ($\times$3.65) is quite close to the "true" average of positive

Table 11: Coefficient Estimates and Shapley Values (Heterogeneous Effects)

	(1)	(2)	(3)	(4)	(5)
		First Step			Second Stage
D_1 INV	-0.787***				
	(0.15)				
PROV w/o INV	1.295***				
	(0.06)				
D_2 EXP		-0.174			
		(0.32)			
PROV w/o EXP		1.212***			
		(0.16)			
D_3 CPP			2.337***		
			(0.15)		
PROV w/o CPP			0.254***		
			(0.06)		
D_4 TBT				1.905***	
				(0.38)	
PROV w/o TBT				0.519***	
				(0.10)	
$SPROV^+$					1.295***
					(0.06)
$SPROV^-$					-0.787***
					(0.15)
$SV(D_n)$	-0.525	0.045	1.558	0.854	
$\exp SV(D_n)$	$\times$ 0.59	$\times$ 1.05	$\times$ 4.75	$\times$ 2.35	
$\exp \beta_n$	$\times$ 0.46	$\times$ 0.84	$\times$ 10.35	$\times$ 6.72	
$\exp \beta^+$					$\times$ 3.65
$\exp \beta^-$					$\times$ 0.46

Notes: Standard errors are in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

provisions ($\times$4.00).[24] Moreover, the estimated effect of the lone negative provision INV ($\times$0.46) is also very close to its "true" average effect ($\times$0.5).

In the next section, we provide a more robust simulation analysis (akin to the setup in this section) but instead for a cross-section of our *full dataset*.

5.4 Simulations on a Full Dataset

We now perform simulations on a cross-section of our full dataset. We have constructed theoretically-driven fictitious trade according to the following algorithm:

$$\begin{aligned}\ln SX_{ij} &= \left(\sum_{n=1}^{N/2}(2+e_{1,ij})D^{+}_{n,ij} + \sum_{n=N/2+1}^{N}(1/2+e_{2,ij})D^{-}_{n,ij} + e_{3,i}\lambda_i + e_{4,j}\lambda_j - 0.7\ln Dist_{ij}\right)\\ &\quad \times(100+e_5), \qquad (10)\end{aligned}$$

where SX_{ij} denotes a simulated value of X_{ij}, λ_i and λ_j are dummies (explained below) that take values of 1 or 0 for countries i and j, respectively, $Dist_{ij}$ is the distance in kilometers between countries i and j, and $e_u \sim N(0,1), \forall u \in \{1,..5\}$. We have randomly assigned half of the substantive provisions as positive ($D^{+}_{n,ij}$) and half as negative ($D^{-}_{n,ij}$). Each positive provision doubles the level of trade between trading partners and each negative provision halves trade. There are $N = 164$ substantive provisions and 138 countries in our dataset. The provisions are the actual provisions signed by the countries by the year 2017.

The simulated trade equation (10) is based on trade theory because, along with the policy variables that affect trade (i.e., provisions), it allows for countries' GDPs and multilateral resistance terms (controlled for by λ_i and λ_j) and distance costs ($Dist_{ij}$).

We have estimated the simulated trade data in multiple ways. The first method (labeled "All Individual") includes a dummy for each provision on the right-hand side. The second method uses the Shapley technique described earlier (labeled "Shapley Individual"). In each method, we have created two variables, grouping the sum of provisions according to the coefficient sign. We then estimated the effect of these count variables on trade, for which estimates are reported in Table 12. We know that by the construction of equation (10), the coefficients should be close to 0.69 for the positive provisions and -0.69 for the negative provisions.

The second stage positive and negative provisions' effects using the first method (for stage 1), reported in column 1 of Table 12, are -0.176 for positive provisions and 0.245 for negative provisions. This method categorized correctly only 39% of the provisions, consistent with signs contradicting expectations.

[24] $4 = (2 \cdot 4 \cdot 8)^{\frac{1}{3}}$

By contrast, the Shapley Individual method reported in column 2 of Table 12 was able to categorize 67% of the provisions correctly and the signs of provisions are as expected, positive for the positive provisions and negative for the negative ones. However, the incorrect categorization of nearly one-third of the provisions has an impact on the magnitude of the coefficient estimates, which show attenuation bias.

Table 12: Estimates with Simulated Trade Data on a Full Dataset

	(1) All Individual	(2) Shapley Individual	(3) All Grouped	(4) Shapley Grouped
Positive Provisions	-0.1762***	0.2767***	0.5046***	0.5107***
	(0.06)	(0.01)	(0.05)	(0.05)
Negative Provisions	0.2415***	-0.1338***	-0.4405***	-0.4201***
	(0.07)	(0.01)	(0.04)	(0.04)
Distance	-0.8851***	-0.5842***	-0.7319***	-0.6808***
	(0.07)	(0.07)	(0.08)	(0.08)
Observations	16628	16628	17074	17074
OriginFE	Yes	Yes	Yes	Yes
DestinationFE	Yes	Yes	Yes	Yes
% Correctly Predicted	39.024	67.073	63.636	90.909

Notes: Estimation uses OLS.

OLS Robust standard errors are in parentheses, clustered by country pair.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The Shapley Value (defined in equation (8)) can be generalized to capture the sign of a group of provisions, rather than a particular provision. By imposing a common coefficient (or average treatment effect) on a group of provisions, we reduce the “dimension” of the problem and aid in reducing the inflated values of estimates of dummy variables.

Therefore, we performed two additional simulations. Instead of assigning provisions randomly in the simulated data, we assigned all substantive “liberalizations” as positive provisions and all substantive “obligations” as negative provisions. The purpose of this is to reduce the “dimensionality” of the regressions. Then, instead of estimating the effect of each individual provision, we have grouped them into their policy areas by adding all substantive provisions in each policy area by their classification (liberalization vs. obligation). In total, we have reduced the dimensionality from estimating 164 coefficients of individual substantive provisions to estimating coefficients for only 20 policy groups of liberalizations (positive provisions using SVs) and obligations (negative provisions using SVs).

Column 3 of Table 12 reports the results when we estimate all *groups* individually; these are referred to as “All Grouped.” This method mimics the one used in column 1 for all individual provisions. Now, we observe the correct signs in the estimates of the positive and negative provisions. However, the success rate is nearly 64%. Although this method is successful in providing reasonably accurate average estimates for the effects of the provisions (which should be 0.7 and -0.7), it fails in identifying which provisions (or their groups) are, in fact, positive or negative.

With the "Shapley Grouped" method, we can still obtain reasonably accurate estimates of the average treatment effects of the positive provisions and of the negative provisions, but nearly perfect categorization of the effects of the provision groups. Applying the Shapley method to the grouped data has a success rate of 91%, as highlighted in column 4 of Table 12.

Several interesting characteristics surface from observing the distribution of the coefficients obtained in the first stage that was used to assign each provision to a positive or negative group, reported in Figure 1. The coefficients of the provisions estimating all individual provisions, shown in Figure 1a, have a very high variance. Some of the coefficients are as high as 40 or as low as -40, revealing a huge bias away from their "true" effects. The distribution of the Shapley Individual method, depicted in Figure 1b, has a much lower variance, but some of the individual coefficients are as high as 5 and as low as -5, and some are still incorrectly categorized. The green distribution should be *almost entirely* positive and the red distribution should be almost entirely negative.

The *grouped* results shown in the bottom figures are much better. In particular, focus on Figure 1d. The distribution of the Shapley Grouped method in Figure 1d reveals that the positive and negative groups are correctly categorized with only *slight overlap* in their respective distributions, as expected. In fact, *only two* groups of negative provisions were incorrectly categorized as positive.

6 Econometric Specifications and Empirical Results

6.1 Methodology for Econometric Specifications

As discussed in section 3, the modern theoretical foundations for bilateral trade flow and bilateral FDI stock determinants suggest that such flows/stocks are explained well using structural gravity equations, cf., Bergstrand and Egger (2007), Ramondo and Rodriguez-Clare (2013), Arkolakis et al. (2018), and Anderson et al. (2019).[25] For instance, in the context of the model in Anderson et al. (2019), the determinants of the bilateral FDI from

[25] **Referee Note:** Should our reference to these four studies be sufficient, our paper can be shortened by eliminating section 3.

origin country i to affiliates in destination country j in year t (FDI_{ijt}) is the gravity equation:

$$FDI_{ijt} = \frac{\beta\phi^2\eta_i^2}{1-\beta+\beta\delta_{j,M}}\omega_{ijt}\frac{E_{it}}{P_{it}}\frac{Y_{jt}}{M_{it}}\epsilon_{ijt}^F, \tag{11}$$

if $FDI_{ijt} = \omega_{ijt}M_{it} > 1$ (and 0 otherwise), where E_{it} is total expenditures in country i (on consumption goods, physical capital investments, and technology capital investments) in year t, P_{it} is a multilateral index of prices in country i in year t on all types of goods, Y_{jt} is a measure of national output in j in year t, M_{it} is the technology capital stock in i in year t, ω_{ijt} is a measure of (policy and non-policy) openness of country j to country i's technology capital, β is the standard time-discount factor, ϕ is the (Cobb-Douglas) share of the global technology capital stock used in production of output, η_i is the share of country i's technology capital as a share of country j's global technology capital stock, and $\delta_{j,M}$ is the technology capital "adjustment cost," analogous to the standard physical capital adjustment costs (in the physical capital accumulation literature).

In the context of their same model, Anderson et al. (2019) also provide a structural gravity

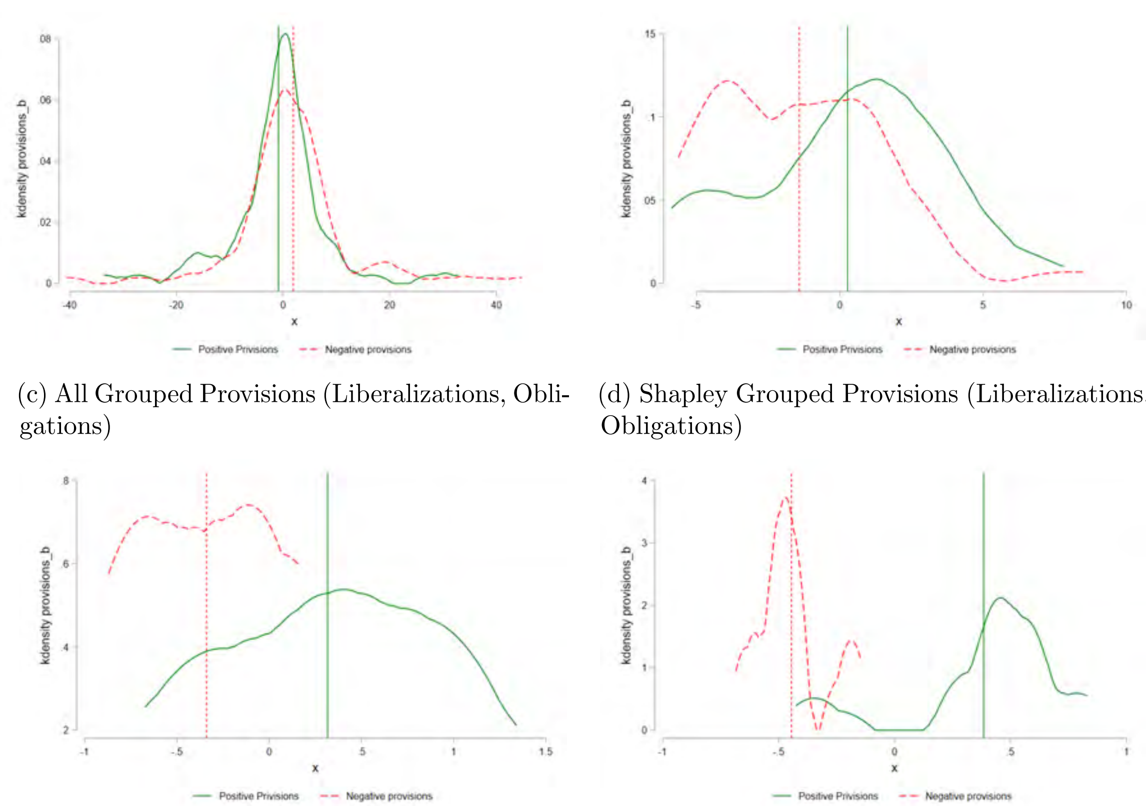
Figure 1: Coefficient Distribution for Simulations using Shapley Method vs. All Provisions

(a) All Individual Provisions

(b) Shapley Individual Provisions

(c) All Grouped Provisions (Liberalizations, Obligations)

(d) Shapley Grouped Provisions (Liberalizations, Obligations)



equation for trade from origin country i to destination country j in year t ($TRADE_{ijt}$):

$$TRADE_{ijt} = \frac{Y_{it}E_{jt}}{Y_t^W}\left(\frac{t_{ijt}}{\Pi_{it}P_{jt}}\right)^{1-\sigma}\epsilon_{ijt}^T, \tag{12}$$

where Y_t^W is world GDP in year t, Π_{it} is the outward multilateral price index of country j in year t, and σ is the elasticity of substitution in consumption.

As standard, we employ PPML estimation using three-way (it, jt, and ij) fixed effects. We include intra-national trade *and* intra-national (or domestic) investment; an additional benefit of the MREID data set is that we have observations on intra-national, or domestic, investment (e.g., iit).

As detailed in section 5.3, an initial set of regressions determines an estimate of β_{SP} for either bilateral trade or FDI. In particular, we estimate:

$$X_{ijt} = \exp\left(\beta_{SP}(SPROV_{ijt} \times BRDR_{ij}) + \lambda_{it} + \lambda_{jt} + \lambda_{ij} + \chi_{ijt}\right) \times \epsilon_{ijt}, \tag{13}$$

where X_{ijt} can be bilateral trade or FDI (with the latter measured here as the number of country i's GUOs' affiliates in country j in year t), BRDR is a dummy that takes the value of 1 if the flow is international ($i \neq j$) and 0 if domestic ($i = j$), and χ_{ijt} represents any other relevant time-varying bilateral determinants. As noted by Heid et al. (2021), our international dummy (BRDR) ensures independence from any specific country choice and avoids systematic variations with DTAs. This equation includes a full set of time-varying origin-country, time-varying destination-country, and country-pair fixed effects (denoted by λ). Finally, the equation includes other covariates (χ_{ijt}) that might affect trade or FDI; specifically, we include dummies for the presence or absence of (shallow) FTAs and of BITs, which do not have DTA provisions.

After this initial step, we estimate the structural gravity equivalent to equation (7)[26]:

$$X_{ijt} = \exp\left(\begin{array}{c}\left(\ \beta_n \cdot D_{n,ijt} + \beta_{SP-n} \cdot (SPROV_{ijt} - D_{n,ijt})\ \right) \times BRDR_{ij} \\ +\lambda_{it} + \lambda_{jt} + \lambda_{ij} + \chi_{ijt}\end{array}\right) \times \epsilon_{ijt}. \tag{14}$$

We repeat this process 164 times (for each substantive provision). Then we use the estimated coefficients of equations (13) and (14) to obtain the Shapley Values, employing equation (8). Next, we classify each provision as positive ($SPROV^+$) or negative ($SPROV^-$) according to its Shapley Value.

[26] We include domestic data and interact all variables with the domestic border dummy BRDR.

In a subsequent step, we estimate the following equation:

$$X_{ijt} = \exp\left(\begin{array}{c}\left(\underbrace{\beta^{+}_{SP}SPROV^{+}}_{SV(D_n)>0} + \underbrace{\beta^{-}_{SP}SPROV^{-}}_{SV(D_n)<0}\right) \times BRDR_{ij} \\ +\lambda_{it} + \lambda_{jt} + \lambda_{ij} + \chi_{ijt}\end{array}\right) \times \epsilon_{ijt}, \tag{15}$$

where $SPROV^{+}$ is the sum of all provisions with a positive SV in the first step and $SPROV^{-}$ is the sum of all provisions with a negative SV in the first step.

A standard concern in gravity settings is that country-pair-specific pre-trends drive both trade and the decision to engage in a trade agreement, creating potential endogeneity bias. The standard setting in Baier and Bergstrand (2007) may not fully address this issue as it only differences out country-pair-specific means. We address this issue by using a triple difference estimator, the difference between two difference-in-differences estimators; the interaction of $SPROV^{+}$ and $SPROV^{-}$ with the BRDR dummy is effectively a triple difference. In this scenario, Olden and Møen (2022) posit that the triple difference estimator remains unbiased if the bias is uniform in both estimators. The rationale behind this is that the disparity between two biased difference-in-differences estimators remains unbiased, given that the bias is consistent across both estimators. This condition necessitates only the adherence of a single parallel trend assumption.

Note that in the first stage we use another additional difference (between equations (13) and (14)) to estimate the SVs, rendering a quasi quadruple difference estimator for the individual provision estimates. Furthermore, and importantly, we use the SVs obtained using the (ITDP-E) *trade* dataset to estimate effects *on FDI*. Conversely, we use the SVs obtained using the (MREID) *FDI* dataset to estimate effects *on trade*. This ameliorates reverse causality concerns because the datasets are independent.

6.2 Empirical Results

6.2.1 Benchmark Results

Table 13 reports the benchmark effect of a dummy variable for the presence or absence of DTAs on bilateral trade flows in Panel A and FDI in Panel B. The estimates of the DTA dummy reported in column 1 (3) of Table 13 suggest that country pairs with a DTA exhibit 29%[27] (6%) higher trade flows (FDI) than those without a DTA.[28] However, these estimates may be biased because column 1 (3) results do not include controls for other types of agreements, such as country-pairs with shallow FTAs and with BITs. In columns 2 and 4,

[27] $\exp(0.2577) = 0.29$

[28] Since domestic trade (investment) is included, the effects of DTAs are on international trade (investment) relative to domestic trade (investment).

we additionally include controls for shallow FTAs and BITs; this lowers non-trivially the estimate in column 2 relative to column 1 in Table 13. However, this lowering in Panel A makes economic sense as the FTA control accounts for tariff liberalizations while the DTA dummy in column 2 now accounts for non-tariff factors.

Several interesting insights surface when focusing on the benchmark effects of DTAs on FDI (measured by the number of affiliates) in Panel B of Table 13. DTAs have a positive and significant effect on FDI, albeit at a lower magnitude than on trade. The estimated coefficient of the DTA dummy on FDI reported in column 3 of Table 13 suggests that country-pairs with DTAs exhibit 6% more foreign affiliates than those without DTAs. The effect after controlling for shallow agreements (FTAs or BIT) is quantitatively similar. BITs have an insignificant effect on FDI.

Moreover, a notably interesting outcome from these benchmark dummy-variable results is the potential implication that deep "trade" agreements – on average – *increase* FDI, which is consistent with the interpretation that FDI serves as an "export platform"; trade and FDI are essentially *complements* with respect for DTAs' provisions, cf., Antràs et al. (2024). Accordingly, the net effect of liberalizations and obligations is to increase FDI, increase the numbers of affiliates abroad, and facilitate the "global value chain." However, we will show later that – using our Shapley Value methodology – trade and FDI are essentially *substitutes* with respect to DTAs' provisions.

Table 13: Effects of DTA Dummy on Trade and FDI, Benchmark

	Panel A: Trade		Panel B: FDI (affiliates)	
	(1)	(2)	(3)	(4)
DTA Dummy	0.2577***	0.1060**	0.0503*	0.0561*
	(0.08)	(0.05)	(0.03)	(0.03)
Observations	190440	190440	190440	190440
Pseudo R^2	0.998	0.998	0.998	0.998
Origin-Year FE	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes
FTA/BIT controls	No	Yes	No	Yes

Notes: Estimation uses PPML. pair.
Robust standard errors are in parentheses, clustered by country pair.
Domestic data are included. FDI is measured as number of affiliates.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 14, we report in the first row the results using the sum of the individual provisions within the DTAs ($SPROV_{ijt}$), instead of the dummy variables. The coefficient estimates reflect the marginal effect of one more provision. The results are *qualitatively* similar to those of the DTA dummies reported in Table 13; recall from results reported in section 5 that a DTA dummy is capturing multiple provisions. By contrast, the coefficients here reflect the marginal effect of a single provision (substantive or non-substantive). The marginal effect of a provision within each DTA has a positive and significant effect on bilateral trade *and* FDI, with a stronger effect on the former than the latter. The result reported in column 1 of Table 14 suggests that including one more individual provision in a DTA increases trade by 0.06%. The effect on FDI of including one more individual provision in a DTA, reported in column 4 of Table 14, is only one-sixth of the former, 0.01%. Nevertheless, we emphasize that these are (average) estimates of a randomly selected provision; in reality, some provisions should increase trade (FDI) and some should decrease trade (FDI). Estimates of such likely *heterogeneous* effects would be useful to policy makers and firms' managements.

Table 14: Effects of Provisions on Trade and FDI

	Panel A: Trade			Panel B: FDI (affiliates)		
	(1)	(2)	(3)	(4)	(5)	(6)
DTA Provisions	0.0006*			0.0001**		
	(0.00)			(0.00)		
DTA Substantive Provisions		0.0032*			0.0006**	
		(0.00)			(0.00)	
DTA Non-Substantive Provisions			0.0008*			0.0002**
			(0.00)			(0.00)
Observations	190440	190440	190440	190440	190440	190440
Pseudo R^2	0.998	0.998	0.998	0.998	0.998	0.998
Origin-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes	Yes	Yes
FTA/BIT controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimation uses PPML. Robust standard errors are in parentheses, clustered by country pair. Domestic data are included. FDI is measured by number of affiliates.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In Table 14, we find that the effects of including *substantive* provisions in DTAs have higher impacts on trade and FDI than increasing any random provision. The result reported in column 2 of Table 14 suggests that including one additional substantive provision increases trade by 0.32%. The effect of including one additional substantive provision on FDI reported in column 5 of Table 14 is lower (0.06%) than that on trade, but consistent with previous estimates in Larch and Yotov (2025) and Osnago et al. (2019).

Conversely, non-substantive provisions in DTAs have much lower effects on trade and FDI than substantive provisions; they are four times lower for trade and three times lower for FDI. In what follows, we focus on substantive provisions.

A useful way to interpret the results reported in Table 14 is to depict the (average) conditional marginal effects of the sum of substantive provisions depicted in Figure 2. For example, in Figure 2a, we observe that a DTA composed of 60 substantive provisions increases trade by 21% on average (i.e., exp(0.192)) and a DTA composed of 120 substantive provisions increases trade by around 40%.[29] The same number of substantive provisions (60 and 120) increases FDI by 4% and 7%, respectively, as shown in Figure 2b.

Table 15 replicates the specifications of Table 14, but now using the log of the sum of substantive provisions.[30] The coefficients reported in Panel A (B) of Table 15 are the trade (FDI) *elasticities* of DTA substantive provisions. For example, focusing on the substantive provisions reported in columns 2 and 5 of Table 15, increasing by 1% the number of substantive provisions increases trade by 6% and FDI by 1%, respectively. Overall, the results align with the linear sum reported in Table 14.

Table 15: Log of Provisions

	Panel A: Trade			Panel B: FDI (affiliates)		
	(1)	(2)	(3)	(4)	(5)	(6)
Log Provisions	0.0370**			0.0090**		
	(0.02)			(0.00)		
Log Substantive Provisions		0.0608**			0.0121**	
		(0.03)			(0.01)	
Log Non-Substantive Provisions			0.0381**			0.0093**
			(0.02)			(0.00)
Observations	190440	190440	190440	190440	190440	190440
Pseudo R^2	0.998	0.998	0.998	0.998	0.998	0.998
Origin-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes	Yes	Yes
BIT/FTA control	Yes	Yes	Yes	Yes	Yes	Yes

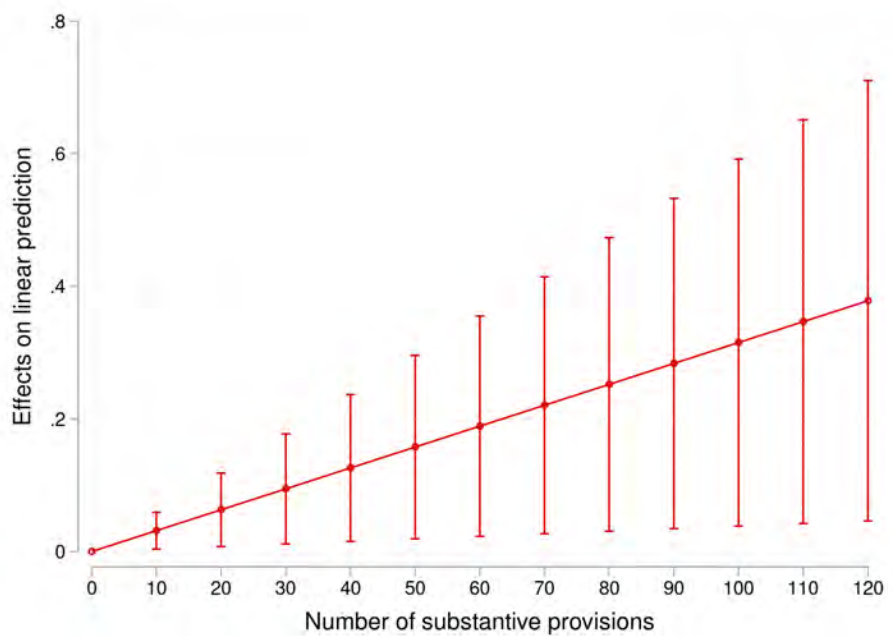
Notes: Estimation uses PPML. Robust standard errors are in parentheses, clustered by country pair, Domestic data are included. FDI is measured as the number of affiliates.
Log stands for inverse hyperbolic sine (compatible with zeros)
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[29] Note that the typical DTA includes between 20 and 50 provisions; hence, the results in Figure 2a are consistent with those in Table 13. The bars represent 95% confidence intervals.

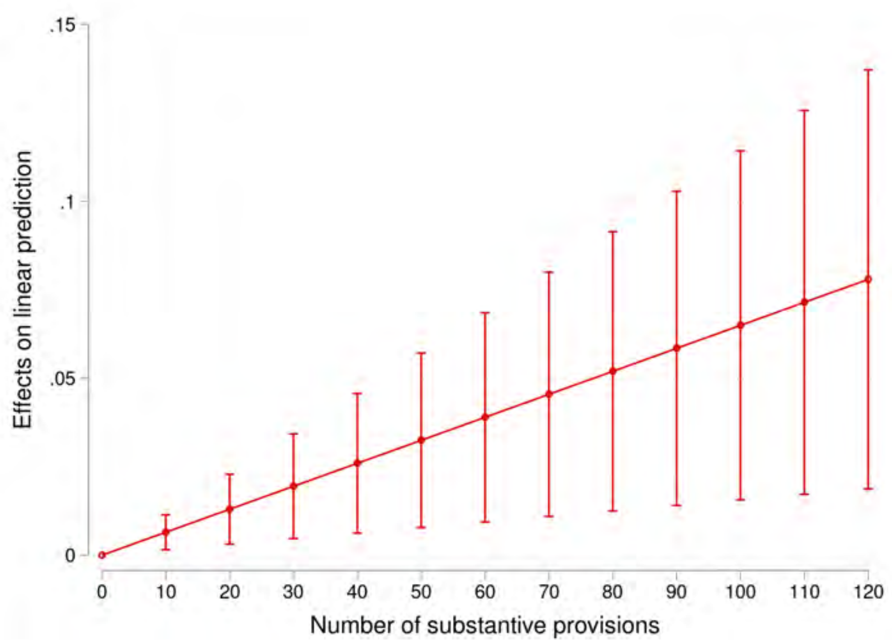
[30] We have taken the inverse hyperbolic sine of provisions, which is compatible with using zeros and has a very similar value and interpretation to the log.

Figure 2: Conditional Marginal Effects of the Sum of Substantive Provisions

(a) Trade



(b) FDI (Affiliates)



Notes: 90% confidence intervals

Table 16 reports the results of non-linear effects of the number of substantive provisions on trade and FDI. The results shown in column 1 of Table 16 suggest a strong quadratic relationship between the number of substantive provisions and trade. However, we do not observe any significant non-linear relationship between the sum of substantive provisions and FDI.

Table 16: Substantive Provisions' Non-linear Effects

	(1) trade	(2) FDI (affiliates)
Substantive Provisions	0.0135***	-0.00004
	(0.00)	(0.00)
Substantive Provisions Squared	-0.0001***	0.00001
	(0.00)	(0.00)
Observations	190440	190440
Pseudo R^2	0.998	0.998
Origin-YearFE	Yes	Yes
Destination-YearFE	Yes	Yes
PairFE	Yes	Yes
FTA/BIT control	Yes	Yes

Notes: Estimation uses PPML. Robust standard errors are in parenthesis, clustered by country pair.
Domestic data are included. FDI is measured as number of affiliates.
Log stands for inverse hyperbolic sine.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

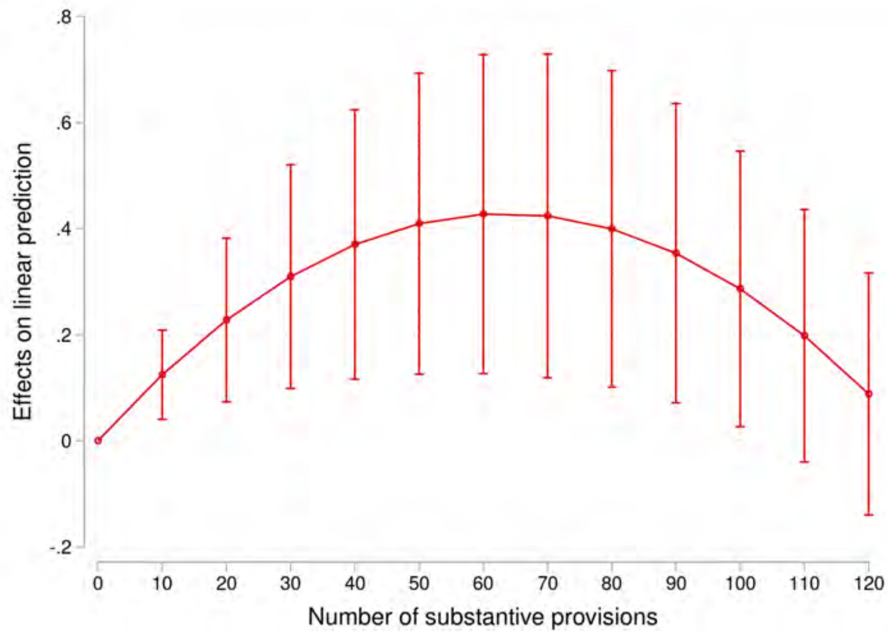
Interpreting the non-linear results of Table 16 is easier when we plot the (average) conditional effects in Figure 3. The linear coefficient in column 1 of Table 16 was positive and significant. Thus, the trade effect of substantive provisions increases initially by increasing the number of substantive provisions in DTAs, as seen in Figure 3a. However, after a turning point of around 60 substantive provisions, increasing the number of provisions has an increasingly lower effect since the quadratic coefficient estimate reported in column 1 of Table 16 was negative and significant.[31] The (average) conditional marginal effects of the non-linear specification of substantive provisions on FDI are shown in Figure 3b. As expected from the non-significant results reported in column 2 of Table 16, we observe only a mild upward trend in these results (with statistical significance only at 120 provisions).

An intuitive interpretation of the non-linear effect of the number of provisions on trade likely reflects *heterogeneity* in the effects of substantive provisions on trade. If some provisions harm trade, we can easily understand the effects exhibited in Figure 3a. For example, a DTA with zero provisions has a null effect on trade, as expected. However, a DTA with 60

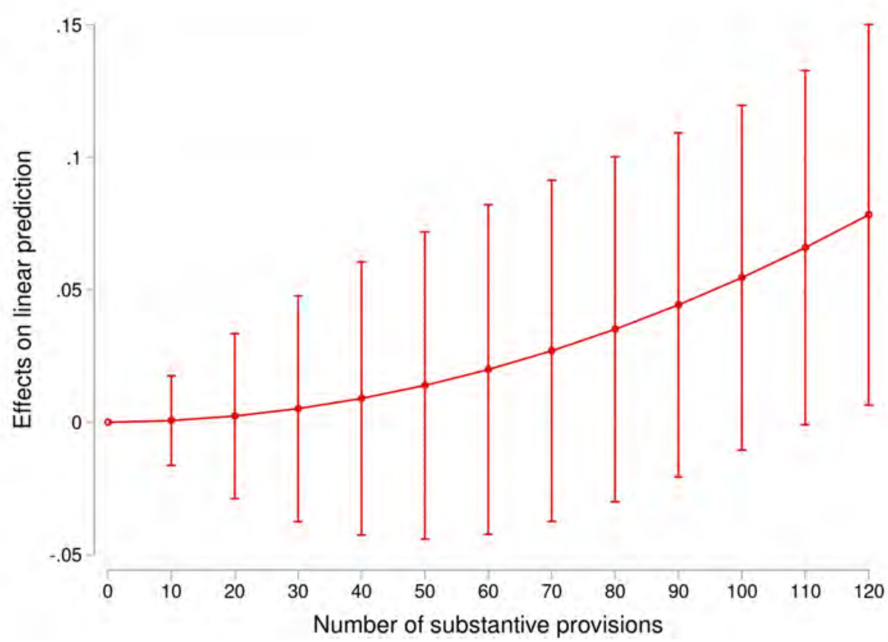
[31] The precise turning point is 67.5, calculated by: $-\frac{0.0135}{2\times 0.0001}$

Figure 3: Conditional Marginal Effects of the Number of Substantive Provisions

(a) Trade



(b) FDI (Affiliates)



Notes: 90% confidence intervals

positive and 60 negative provisions might also have a null effect on trade. This can explain why the average marginal effect of a DTA with 120 provisions is insignificantly different from zero in Figure 3a.

6.2.2 Positive and Negative Provisions (First Stage Results)

This section summarizes the results of the first stage of estimation, estimates of the Shapley Values (SVs) of substantive provisions. After running the regressions in equations (13) and (14), each provision is categorized as positive-trade (negative-trade) if the SV is positive (negative) in the trade regression and positive-FDI (negative-FDI) if the SV is positive (negative) in the FDI regression. This is performed with two methods: the Individual SV and Grouped SV methods. The Individual SV estimates an SV for each of the 164 substantive provisions. The Grouped SV method estimates an SV for each group of policy area and provision type (liberalization/obligation).

The results of the first stage are summarized in Table 17.[32] The Individual SV method categorizes 79 provisions (48%) as positive-trade and 85 provisions (52%) as negative-trade, cf., Panel I. Using the same method, 97 provisions (59%) are categorized as positive-FDI and 67 provisions (41%) as negative-FDI. This is our first evidence that there are *considerable fractions* of substantive provisions that negatively affect trade and FDI.

The first-stage SVs are crude estimates of the quantitative effects of provisions.[33] Therefore, we should expect positive SVs that are higher than the average marginal effect of all (substantive) provisions, which we estimated earlier (0.0032; see Table 14). The average SV for positive-trade is 0.0677, confirming our intuition. Similarly, the SV for positive-FDI, 0.0043, is higher than the estimated effect for all substantive provisions, 0.0006; see Table 14. This also aligns with our theory.

Table 17: Descriptive Statistics of Shapley Values (First Stage)

	Panel I: Individual SVs			Panel II: Grouped SVs		
	Number	Mean	S.D.	Number	Mean	S.D.
Positive-Trade SV	79	0.0677	0.0621	100	0.0665	0.0790
Negative-Trade SV	85	-0.0726	0.0487	64	-0.0784	0.0540
Positive-FDI SV	97	0.00426	0.00599	82	0.0102	0.00751
Negative-FDI SV	67	-0.00696	0.00583	82	-0.00886	0.00603

The Grouped Shapley method, which showed higher performance in characterizing provisions in our simulations, show very similar SVs for trade in terms of magnitudes – at least in terms of average effects, cf., Panel II in Table 17. This method yields 61%

[32]The details of the Shapley Values for each provision can be found in the Online Supplement.
[33]In the next section, we provide more precise estimates in the second stage.

of the provisions (or 100 in total) in the positive-trade group with the remainder in the negative-trade group. For FDI, the Grouped SVs are larger on average than those for the Individual SV case and the provisions are evenly split.

The Grouped Shapley method consisted of estimating a SV for each group of provisions; therefore, this method imposes a common coefficient for all provisions within each group. Each provision was grouped according to its policy area *and* its type in terms of liberalization or obligation (according to the World Bank). Four out of the 13 policy areas with substantive provisions (Subsidies, Migration (Visa & Asylum), State-Owned Enterprises, and Competition Policy) had only obligation provisions with no liberalizations. This left us with a total of 22 groups of provisions, 9 liberalization and 13 obligation.

Table 18 reports the groups, the averages of the Individual SVs (by group), and the Grouped Shapley Values of each group. We note several results.

Table 18: Shapley Values by Type and Policy Groups/Areas

		Shapley Value (Indv)		Shapley Value (Group)	
Type	Policy Area	Trade	FDI	Trade	FDI
Liberalization	Public Procurement (3)	-0.097	0.006	-0.158	0.011
Liberalization	Investment (11)	-0.073	-0.006	-0.147	-0.011
Liberalization	Sanitary (4)	-0.041	0.006	-0.093	0.021
Liberalization	Export Taxes (14)	-0.011	-0.000	0.001	-0.013
Liberalization	Trade Facilitation (4)	-0.007	0.002	-0.017	0.005
Liberalization	IPR (19)	-0.002	0.001	0.024	-0.000
Liberalization	Tech Barriers (3)	0.005	0.000	0.005	0.000
Liberalization	Rules of Origin (11)	0.039	0.000	0.160	0.003
Liberalization	Services (8)	0.111	-0.006	0.220	-0.009
Obligation	Public Procurement (1)	-0.174	0.012	-0.174	0.012
Obligation	Trade Facilitation (2)	-0.089	-0.003	-0.095	-0.003
Obligation	Investment (2)	-0.085	-0.005	-0.098	-0.005
Obligation	Export Taxes (2)	-0.042	0.004	-0.096	0.004
Obligation	Rules of Origin (3)	-0.031	-0.001	0.037	-0.006
Obligation	IPR (19)	-0.031	0.004	-0.086	0.015
Obligation	Sanitary (16)	-0.017	0.002	0.001	0.020
Obligation	Subsidies (3)	-0.015	-0.003	0.028	-0.003
Obligation	Services (10)	-0.005	-0.006	0.031	-0.017
Obligation	Tech Barriers (16)	-0.004	0.000	-0.006	0.002
Obligation	Migration (3)	0.025	0.002	0.099	0.006
Obligation	State Owned Enter. (8)	0.089	-0.007	0.189	-0.015
Obligation	Competition Policy (2)	0.124	-0.004	0.155	-0.006

Notes: The Individual Shapley Values are the averages per type and policy area group.
The number in parentheses indicates the number of substantive provisions in each group.

First, we note that – using either Individual SVs or Grouped SVs – the World Bank's (so-called) liberalizations and obligations each have a mix of positive and negative trade impacts and positive and negative FDI impacts. Second, for provisions characterized (by the World Bank) as liberalizations, the Individual SVs and Grouped SVs for trade had similar signs (except for Export Taxes and IPR). Analogously, for these liberalization provisions, the Individual SVs and Grouped SVs for FDI had similar signs (except for IPR). It is also noteworthy that some liberalizing policy areas have a negative effect on trade. Since the effec on FDI seems to be positive in most of them, these policies seem to encourage FDI over cross-border trade, substituting trade rather than complementing it.

For provisions characterized as obligations (by the World Bank), the Individual SVs and Grouped SVs for trade had similar signs (except for Rules of Origin, Sanitary, Subsidies, and Services). For these obligation provisions, the Individual SVs and Grouped SVs for FDI had similar signs with no exceptions. This is verified quantitatively in Table 19, where SV Trade(Indv) and SV Trade(Group) have a high correlation coefficient (0.925), and SV FDI(Indv) and SV FDI(Group) have a high correlation coefficient (0.819).

Third, by contrast, the Trade SVs and FDI SVs had *opposite* signs by each group of provisions; this holds whether one uses Individual SV Trade and FDI provisions or Grouped SV Trade and FDI provisions. While difficult to observe using Table 18, Table 19 reports the correlations between SV Trade (Indv) and SV FDI (Indv) and between SV Trade (Group) and SV FDI (Group). For the former, the correlation coefficient is -0.550 and is statistically significant and for the latter the correlation coefficient is -0.421 and is also statistically significant. This is the first evidence suggesting that narrow groups of provisions that positively affect trade negatively affect FDI; this holds whether using the Individual SV method (and averaging provisions' effects) or using the Grouped SV method.

Table 19: Correlation Matrix of Shapley Values

	SV Trade (Indv)	SV FDI (Indv)	SV Trade (Group)
SV FDI (Indv)	-0.550***	1	
SV Trade (Group)	0.925***	-0.527**	1
SV FDI (Group)	-0.370*	0.819***	-0.421*

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The Individual Shapley values are the averages per type and policy area group.

Fourth, some of the groups of provisions that are exceptions to the previous point – that is, the Trade and FDI SVs both have the *same* signs – are consistent with previous results in these areas. Regarding migration, the extant literature has documented the positive impact of migration on FDI and multinationals (e.g., Burchardi et al., 2019; Choi et al., 2024; Cuadros et al., 2019; Glennon, 2023; Morales, 2023) and on trade (e.g., Cuadros et al., 2019; Gould, 1994; Ottaviano et al., 2018; Peri and Requena-Silvente, 2010), consistent with our Obligation-Migration SVs. The model developed by Ornelas and Turner (2024) is compatible with our positive results for trade and FDI related to rules of origin, consistent with our Liberalization-Rules of Origin SVs. Ornelas and Turner (2024) show theoretically that DTAs with provisions related to rules of origin increase foreign investment and induce excessive trade within the trading bloc, explaining our positive result both in trade and FDI for Liberalization-Rules of Origin. Finally, Larch and Yotov (2025) report negative results on FDI of the provisions in the investment policy group, consistent with our negative SVs for Investment provisions and FDI (both Liberalizations and Obligations).

Fifth, when we ignore the World Bank's differentiation between Liberalizations and Obligations and group substantive provisions *only* by their policy categories, we find strong evidence that SVs are *oppositely signed* for trade and FDI. Table 20 reports the mean SVs by the 13 policy areas that contain substantive provisions. Focusing on the Grouped Shapley Values in Table 20, most policy areas exhibit opposite signs of SV for trade and FDI, except for Investment, Export Taxes, and Migration.

Table 20: Shapley Values by Policy Area

Policy Area	Shapley Value (Indv)		Shapley Value (Group)	
	Trade	FDI	Trade	FDI
Public Procurement	-0.136	0.009	-0.166	0.011
Investment	-0.079	-0.006	-0.122	-0.008
Trade Facilitation	-0.048	-0.001	-0.056	0.001
Sanitary	-0.029	0.004	-0.046	0.020
Export Taxes	-0.026	0.002	-0.048	-0.005
IPR	-0.016	0.003	-0.031	0.007
Subsidies	-0.015	-0.003	0.028	-0.003
Tech Barriers	0.000	0.000	-0.001	0.001
Rules of Origin	0.004	-0.000	0.098	-0.002
Migration	0.025	0.002	0.099	0.006
Services	0.053	-0.006	0.125	-0.013
State Owned Enter.	0.089	-0.007	0.189	-0.015
Competition Policy	0.124	-0.004	0.155	-0.006

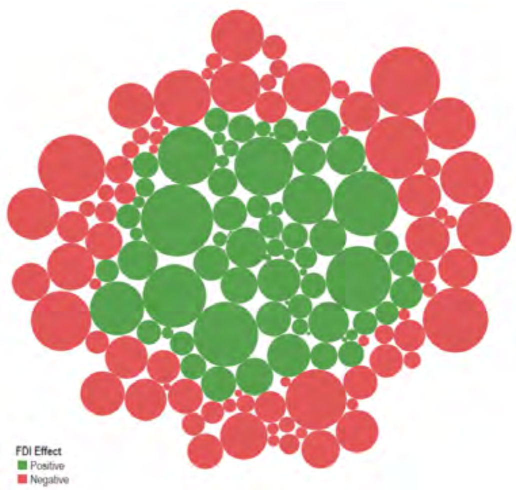
Notes: The second and third columns use averages of Individual Shapley Values.

Figure 4 depicts all 164 provisions in terms of their FDI and trade SVs (proportional to the size of their SV effects, with green for positive and red for negative).

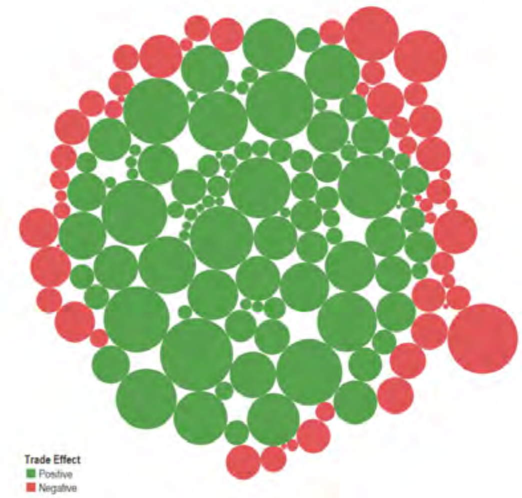
As expected, the Individual SV grouping shown in panels 4a and 4c for trade and FDI, respectively, suggest that policymakers have adopted positive and negative provisions relatively evenly. However, the Grouped SV pie chart for trade shown in panel 4b reveals a different scenario; nearly three-quarters of adopted provisions had a positive SV and only one-quarter of the provisions had a negative SV. The Grouped SV chart for FDI reveals a

Figure 4: Positive and Negative Shapley Values by Provisions

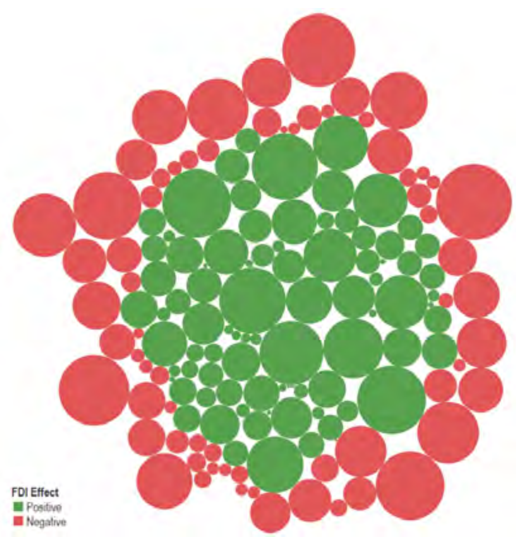
(a) Trade Shapley Values (Indv)



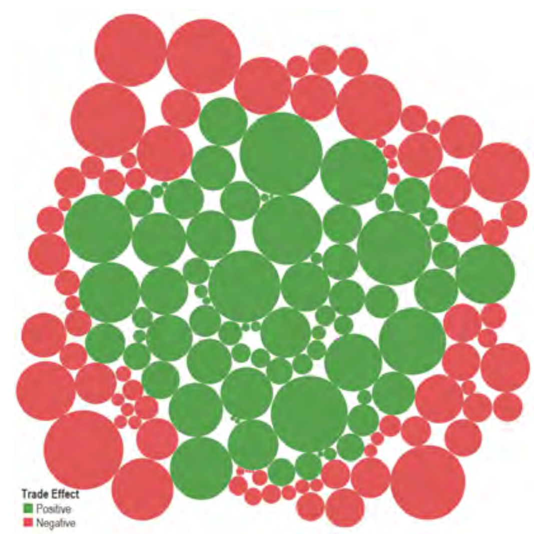
(b) Trade Shapley Values (Group)



(c) FDI Shapley Values (Indv)



(d) FDI Shapley Values (Group)



Notes: The size of the bubble is proportional to the SV of the provision. The color indicates the sign of the Shapley Value.
Source Data: World Bank DTA Database and own elaboration

relatively even split that leans towards more negative than positive FDI provisions in terms of signs.

6.2.3 Effects of Positive and Negative Substantive Provisions (Second Stage Results)

Our theory and the non-linear empirical results reported in Table 16 suggest that provisions have heterogeneous effects on trade and FDI. Those provisions that enhance trade might hinder FDI, and vice versa. Therefore, we applied the Shapley Value method described in Section 5.3 to distinguish positive from negative provisions. In the first step, reported above in Section 6.2.2 and in the Online Supplement, we identified each substantive provision's individual effect's sign (positive or negative) on trade and FDI. Now, in the second step, we create two variables: the first variable adds all the positive substantive provisions (for each of trade and FDI) and the second one adds all the negative substantive provisions (for each of trade and FDI). We will find convincing evidence that trade and FDI are effectively substitutes with respect to DTAs' provisions. For brevity, we focus here on the Grouped Shapley Values approach; similar results using the Individual Shapley Values approach are reported in the robustness analysis in section 7.

Table 21, Panel A, reports the effects on bilateral trade of positive-trade and negative-trade provisions, that is, the sets of provisions that have either positive or negative effects on trade in the first step applying the Grouped Shapley method, using equation (15).[34] As expected and in line with our theory and simulations, positive-trade substantive provisions have a positive and significant effect on trade and negative-trade substantive provisions have a negative and significant effect on trade. However, including one additional positive-trade provision in a DTA increases trade by 2.09%. This is more than *six times* the effect of including one random provision (0.32% in Table 14), as suggested by the typical "count" of provisions approach used in the literature. Furthermore, one additional negative-trade substantive provision in a DTA decreases trade by 3.58%. This is more than *ten times* (in absolute value) the effect of including one random provision. The weighted net effect is close to 0.4% and is not statistically different from the estimate of the effect of the sum of all substantive provisions, 0.32%.[35]

Although the qualitative outcomes in column (1) may not appear surprising, and may be compromised by possible endogeneity, Column 2 of Table 21 reports the effects on bilateral

[34] Recall that with the interaction with the international border dummy, we are using a triple difference estimator that enhances the identification in Baier and Bergstrand (2007) Also, in a robustness analysis later, we will report the results using the Individual Shapley Value method.

[35] Although the marginal effect of negative (grouped) provisions is higher (in absolute value) than the positive ones, negative provisions are only 27% of all substantive provisions. Hence, the overall weighted effect is positive.

trade of *positive-FDI* and *negative-FDI* provisions. That is, those provisions that have positive effects on FDI in the first step turn out in the second step to have *negative* effects on trade. Positive-FDI substantive provisions have a negative and significant effect on *trade* and negative-FDI substantive provisions have a positive and significant effect on *trade*. These results tend to confirm the *countervailing effects* of provisions on trade and FDI, as first noted in Bergstrand and Egger (2007). Yet, we used a completely different FDI dataset in the first step to categorize the provisions, and we obtained highly significant and opposite effects on trade in the second step. This approach has the advantage of minimizing any endogeneity concerns that might arise from using the same dependent variable in the first and second steps, as discussed earlier.

Table 21: Positive and Negative Substantive Provisions (Grouped Shapley Effects)

	Panel A: Trade		Panel B: FDI (affiliates)	
	(1)	(2)	(3)	(4)
Positive-Trade Provisions (Grouped Shapley)	0.0209*** (0.01)		-0.0017 (0.00)	
Negative-Trade Provisions (Grouped Shapley)	-0.0358*** (0.01)		0.0048** (0.00)	
Positive-FDI Provisions (Grouped Shapley)		-0.0144* (0.01)		0.0049*** (0.00)
Negative-FDI Provisions (Grouped Shapley)		0.0121** (0.01)		-0.0017* (0.00)
Observations	190440	190440	190440	190440
Pseudo R^2	0.998	0.998	0.998	0.998
Origin-Year FE	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes
FTA/BIT controls	Yes	Yes	Yes	Yes

Notes: Estimation uses PPML. Robust standard errors are in parentheses, clustered by country pair. Domestic data are included. FDI is measured as number of affiliates.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As further evidence of the apparent result that trade and FDI are effectively *substitutes* with respect to provisions, Panel B of Table 21 repeats the exercise using FDI (measured as the number of affiliates) as the dependent variable. Of note, the coefficient estimates of positive-trade and negative-trade provisions reported in column (3) of Table 21 have negative and positive signs, respectively, and are statistically significant. Note that an additional negative-trade provision has a statistically significant positive effect on FDI of 0.48%, which is larger than a randomly selected substantive provision (0.32%). Column (4) of Table 21 reports the effects of positive-FDI and negative-FDI substantive provisions on the number of affiliates. As expected, these coefficients are positive and negative, respectively, and are statistically significant. Including an additional positive-FDI provision on a DTA increases FDI by 0.49%, one and a half times the effect of increasing a random substantive provision. Including an additional negative-FDI provision on a DTA decreases FDI by 0.17%. The combined *net* effect is 0.18%, which is not statistically different at the 95% confidence level from the effect on FDI that we estimated of the sum of all provisions (0.06%) in Table 14.

A convenient way to summarize the results reported in Table 21 is by plotting in Figure 5 the conditional marginal effects on trade and FDI of the positive and negative SV Groups. We observe that positive-trade and negative-trade substantive provisions have a positive and negative effect on trade, respectively, in Figure 5a and a negative and positive effect on FDI, respectively, in Figure 5b. Analogous interpretations hold for the other two figures, 5c and 5d.

We can use Figure 5 to approximate the effects of DTAs with heterogeneous provisions. For example, focusing on the Grouped Shapley method in Figure 5a, we can obtain the expected effects of a DTA with a particular number of positive-trade and negative-trade provisions on bilateral trade. For instance, suppose a DTA has 70 positive-trade provisions. This would increase trade by 0.7, or roughly 101% (exp(0.7)). But suppose the DTA has 30 negative-trade provisions. This would diminish trade by 0.5, or roughly 39%. The net effect would be an increase in trade of 62%, which is in line with many other estimates.

Since these sets of results are significant and tend to align better with our simulations, in the next section, we employ the Group Shapley approach to study the effects of DTAs on other dimensions of FDI. In particular, we examine the effects of DTAs' provisions on MNEs' *costs, factors of production,* and *revenues.*

6.2.4 MNEs' Costs, Factors of Production, and Revenues

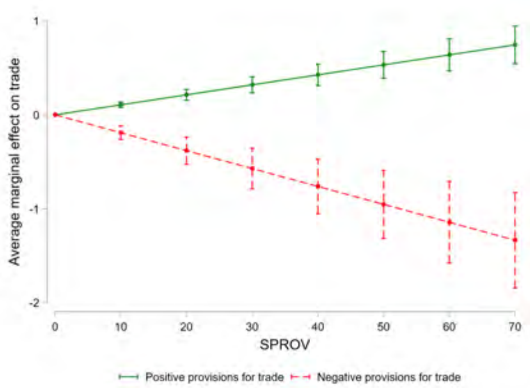
One of the advantages of our new MREID data set is the abundance of variables reported in Orbis on MNEs' "activities." In this section, we examine the potential impacts of Grouped Shapley sets of positive and negative trade and FDI provisions on various *other* economically important MNE variables. The intent is to try to better understand the "mechanism" through which the provisions operate to influence levels of trade and FDI. Specifically, in the context of the theoretical models of trade, FDI, and gravity discussed in section 3 and summarized with our theory-motivated gravity equations in section 6.1, we are interested in whether the provisions likely affect FDI via lowering or raising γ_{ijt}, the measure of bilateral investment costs in section 3 (or raising or lowering ω_{ijt}, the measure of "openness" to FDI discussed in section 6.1).

In this section, we employ the same methodology from above to separate positive and negative provisions, using the number of foreign affiliates in the first step. Table 22 reports the second step, which estimates the effects of the positive-FDI and negative-FDI provisions on each FDI-related variable (Panel I) and the effects of the positive-trade and negative-trade provisions on each FDI-related variable (Panel II).

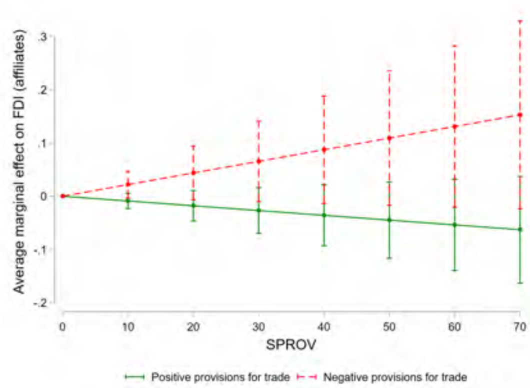
First, and most importantly, the positive-FDI provisions (based on the Grouped Shapley method) *reduce significantly* MNE's costs. The results shown in columns 1 and 2 of Panel I in Table 22 suggest that positive-FDI (negative-FDI) provisions have a negative (positive) and significant effect on total employee costs *and cost per employee*, as expected. In the context of Arkolakis et al. (2018) (discussed here in subsection 3.2 of section 3), a positive-FDI provision (as determined in the first step) has a negative effect on the cost of a country i's

Figure 5: Conditional Marginal Effects of Positive and Negative Substantive Provisions (Grouped Shapley effects)

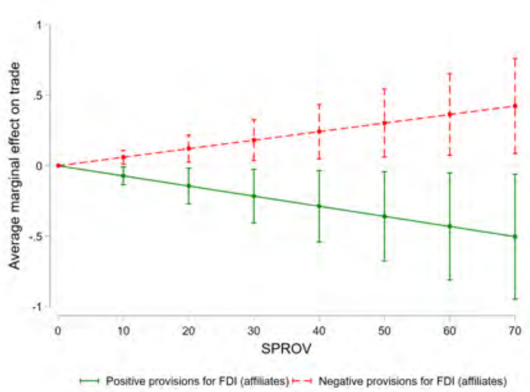
(a) Effects on Trade using SVs of Trade Group Provisions



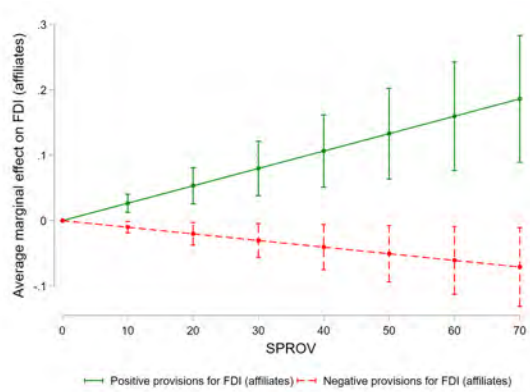
(b) Effects on FDI using SVs of Trade Group Provisions



(c) Effects on Trade using SVs of FDI Group Provisions



(d) Effects on FDI using SVs of FDI Group Provisions



GUO FDI in an affiliate in country l, $\frac{\gamma_{il} w_l^p}{z_l}$, via a reduction in γ_{il}, the *ad valorem* barrier for the GUO in i for FDI in l. This lowers cost per employee, shown in column (2), then lowering overall employee costs, shown in column (1). Note that the quantitative magnitude of one more positive-FDI substantive provision in lowering marginal costs (5%) is approximately the same as that for overall costs, as the effect on the numbers of employees, shown in column (3), is positive as expected but is quite small and is statistically insignificantly different from zero. Conversely, one more negative-FDI provision raises costs per employee as well as overall employee costs, as expected.

Second, as evidence that the results discussed above are not biased by endogeneity, we also provide in Panel II the effects on costs per employee, overall employee costs, and number of employees of (step one's) positive-trade and negative-trade provisions. The effects on costs per employee, overall employee costs, and number of employees have the expected signs. The results reported in columns (1) and (2) of Panel II in Table 22 suggest that positive-trade (negative-trade) provisions increase (decrease) costs per employee and overall employee costs. In other words, those provisions that have a positive impact on trade now increase

Table 22: Positive and Negative Provisions Effects on Various MNE Variables (Grouped Shapley)

	(1) Employee Costs	(2) Costs per employee	(3) Employees	(4) Tangible Assets	(5) Intangible Assets	(6) Revenues
			Panel I			
Positive-FDI Provisions (Grouped Shapley)	-0.0508***	-0.0566**	0.0123	-0.0041	0.0148	-0.0067
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
Negative-FDI Provisions (Grouped Shapley)	0.0287***	0.0382**	-0.0092	0.0018	-0.0135*	0.0055
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
			Panel II			
Positive-Trade Provisions (Grouped Shapley)	0.0436***	0.0106	-0.0109***	-0.0051	-0.0231**	0.0015
	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)
Negative-Trade Provisions (Grouped Shapley)	-0.0859***	-0.0100	0.0187**	0.0087	0.0353*	0.0004
	(0.03)	(0.03)	(0.01)	(0.01)	(0.02)	(0.01)
Observations	18784	16187	190440	190440	190440	190440
Origin-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes	Yes	Yes
FTA/BIT controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimation uses PPML. Robust standard errors are in parentheses, clustered by country pair.
Domestic data are included.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

MNEs' employment costs at affiliates as the number of employees at a foreign affiliate decline. Although we do not observe significant results on cost per employee, the signs of the estimated coefficients are as expected (the opposite from FDI grouped provisions). As shown in column (3), positive-trade (negative-trade) provisions, by lowering (raising) trade costs from country i to country l, lowers (raises) the number of affiliate employees producing in destination country l.

Third, although such provisions have no statistically significant effect on tangible assets of a GUO based in i producing in l as shown in column (4), such provisions have the expected effects – and statistically significant effects – on the GUO's *intangible assets*. One more positive-FDI (negative-FDI) provision increases (decreases) intangible asset investments by approximately 1.5%, though the positive-FDI effect is not statistically significant. Moreover, one additional positive-trade provision *reduces* i's intangible assets investment by 2.3%; this reinforces the interpretation that trade and FDI are effectively substitutes with respect to DTAs. Furthermore, one additional negative-trade provision, by raising trade costs, *increases* intangible assets investments in l by 3.5%. However, we do not find any statistically significant effects of such provisions on revenues of GUOs based in i with affiliates in l.

The results of the positive-FDI and negative-FDI provisions groups (based on the Grouped Shapley method) are summarized in Figure 6.

As expected from the results reported in Panel I of Table 22, Figures 6a and 6b show stark reductions in the employee costs and costs per employees, respectively, associated with positive-FDI provisions. To gain intuition in the results, consider a DTA with 70 positive-FDI provisions. According to the plots, this DTA would reduce the total cost of employees of foreign affiliates by roughly 75% and the cost per employee by around 60%. Naturally, this reduction can be offset by including negative FDI provisions, which increase costs.

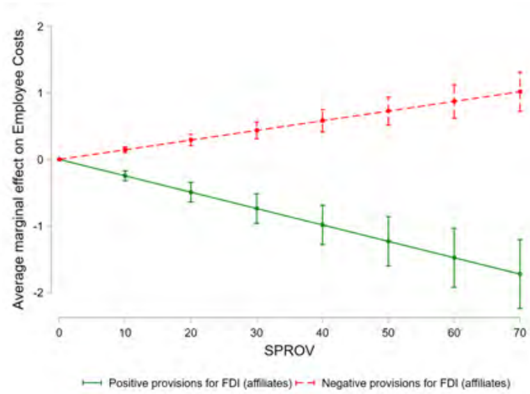
Focusing on effects on the number of employees depicted in Figure 6c, we observe how positive-FDI provisions can have a positive and significant effect on the MNE's affiliate's number of employees, depending upon the number of such provisions. Conversely, negative-FDI provisions can have a negative and significant effect, again depending upon the number of such provisions.

The other factor of production where we observe significant effects is intangible assets, shown in Figure 6f. These results might illuminate why previous studies in this literature found milder and insignificant effects of IPR provisions on FDI (for example, Santacreu, 2021).

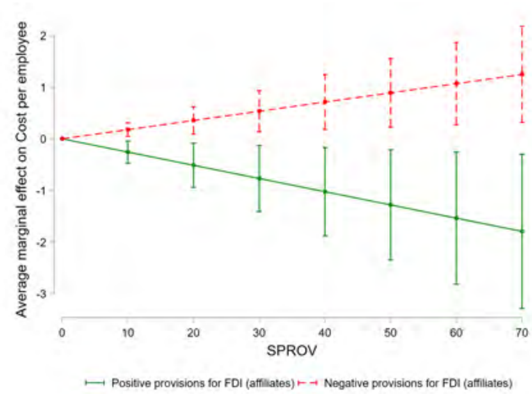
Putting all the pieces together, positive-FDI provisions have a positive effect on increasing some factors of production, notably the numbers of employees and investments in intangible assets. These provisions also reduce the employment costs. Therefore, these firms tend to be more efficient in terms of labor. This is precisely what we observe in Figure 6b when we estimate the effect on the cost per employee (or the ratio of the total costs of employees to the number of employees); positive-FDI provisions reduce the cost per employee. However,

Figure 6: Conditional Marginal Effects of Positive and Negative FDI Provisions on Various MNE Variables (Grouped Shapley)

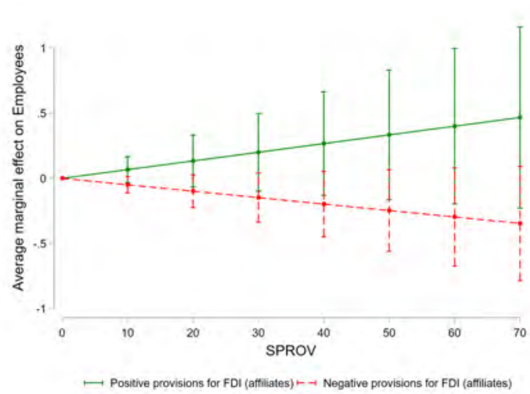
(a) Employee Costs



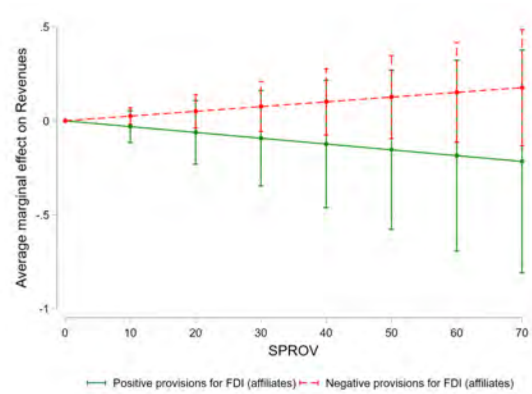
(b) Cost per Employees



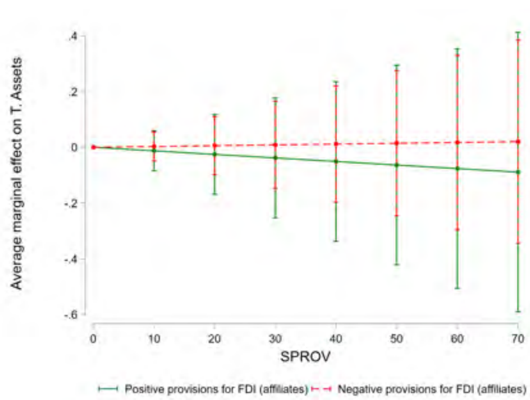
(c) Number of Employees



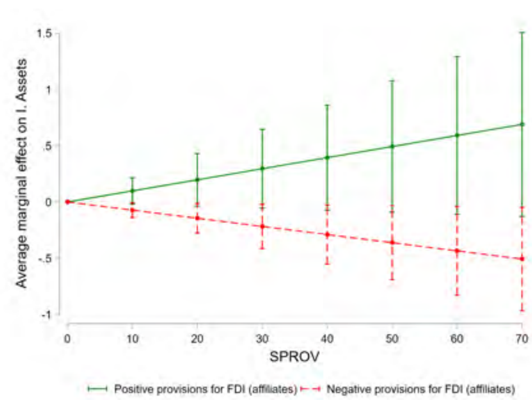
(d) Revenues



(e) Tangible Assets



(f) Intangible Assets



we do not find evidence of positive effects on revenues. In the next section, we provide the results of numerous sensitivity analyses.

7 Robustness Analysis

This section reports the results of several robustness analyses. Section 7.1 reports the results using the Individual Shapley Value method for comparison to the Grouped Shapley Value method. Section 7.2 uses alternative econometric specifications to estimate the effects of provisions on trade and FDI. Section 7.3 examines the effects of positive and negative trade and FDI provisions on costs, factors of production, and revenues *per affiliate*. Section 7.4 reports the estimates of the relative effects of multilateral versus bilateral provisions. Section 7.5 reports estimates of the spillover effects of provisions. The main results are reported in the subsections, and additional results are included in the Online Supplement.

7.1 Individual Shapley Results

In section 6.2.3, we used the Grouped Shapley Value approach to estimate the effects of positive and negative trade and positive and negative FDI provision sets on trade and FDI. For robustness, we provide the analogous results using the Individual Shapley Value method. These results, reported in Table 23, are qualitatively similar to the Grouped Shapley Value method results of Table 21, but there are some minor quantitative differences.

First, as with the Grouped Shapley results, the Individual Shapley results show positive-trade and negative-trade substantive provisions have positive and negative effects on trade, respectively. Importantly, these coefficient estimates are statistically significant. The minor difference is that the Individual Shapley positive-trade and negative-trade coefficient estimates for trade are smaller (in absolute value) than Grouped Shapley ones. By contrast, the Individual Shapley positive-FDI and negative-FDI coefficient estimates for trade are larger (in absolute value) than the Grouped Shapley ones, but are also statistically significant.

Second, for Panel B, neither positive-trade nor negative-trade provisions have statistically significant effects on FDI using the Individual Shapley approach, in contrast to the Grouped Shapley approach where negative-trade provisions have a significant effect on FDI. However, positive-FDI and negative-FDI provisions have similar statistically significant effects on FDI across the two methods.

We also estimated the effects of positive and negative provisions on *costs, factors of production,* and *revenues* using the Individual Shapley method. The results are shown in Table 24. First, overall the results using the Individual Shapley method are similar qualitatively to the Grouped Shapley method results. For positive-FDI and negative-FDI provisions, the partial effects on employee costs, costs per employee, and number of employees have the same signs for both methods, but the effects on employee costs and costs per employee are smaller (in absolute value) for the Individual Shapley method. However, despite similarly sized effects for both methods, the partial effects for number of employees are statistically significant only for the Individual Shapley method. For both methods, negative-FDI provisions have the expected negative effect on investments in intangible assets.

Second, for positive and negative trade provisions on FDI variables, the signs for all coefficients are the same between the two methods. The notable difference is that positive-trade and negative-trade provisions have statistically significant coefficient estimates for employee costs, number of employees, and the size intangible assets using the Grouped Shapley method only.

7.2 Alternative Econometric Specifications

In this section, we consider several alternative estimation techniques and specifications. Some of the results are reported here and, for brevity, some are reported in the Online Supplement.

First, our initial alternative estimation technique was to use OLS rather than PPML. Based upon recent developments, our main estimation technique is PPML to handle depen-

Table 23: Positive and Negative Substantive Provisions (Individual Shapley Method)

	Panel A: Trade		Panel B: FDI (affiliates)	
	(1)	(2)	(3)	(4)
Positive-Trade Provisions (Indv. Shapley)	0.0177***		-0.0008	
	(0.00)		(0.00)	
Negative-Trade Provisions (Indv. Shapley)	-0.0106***		0.0014	
	(0.00)		(0.00)	
Positive-FDI Provisions (Indv. Shapley)		-0.0171***		0.0043***
		(0.01)		(0.00)
Negative-FDI Provisions (Indv. Shapley)		0.0161***		-0.0024**
		(0.00)		(0.00)
Observations	190440	190440	190440	190440
Pseudo $R2$	0.998	0.998	0.998	0.998
Origin-Year FE	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes
FTA/BIT controls	Yes	Yes	Yes	Yes

Notes: Estimation uses PPML. Robust standard errors are in parentheses, clustered by country pair. Domestic data are included. FDI is measured as number of affiliates.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

dent variable zeros and to handle coefficient estimation biases owing to heteroskedasticity, cf., Santos Silva and Tenreyro (2006) and Baier et al. (2017). Table 25, Panels IA and IB provide the respective OLS results to the PPML results in Table 21. First, and most importantly, in a comparison none of the coefficient estimates' signs are different between the PPML and OLS results. Second, although coefficient estimates across the two specifications are identical qualitatively, there are some variances across the specifications' coefficient estimates quantitatively.

Second, Weidner and Zylkin (2021) show that PPML may have biased coefficient estimates in cases like ours, due to the Incidental Parameters Problem (IPP). To address this concern, we used their proposed a bias-corrected PPML estimation technique. The results are shown in Panels IIA and IIB of Table 25. Comparing these results to the comparable results in Table 21 reveals no material differences qualitatively or quantitatively in the results.

Finally, we also estimated several alternative specifications of the main equation. For brevity, the tables for these results are relegated to the Online Supplement, and we simply summarize the results here. First, to further preclude any endogeneity biases we re-estimated the specifications in Tables 21 and 23 using the lagged values of the RHS variables. The

Table 24: Positive and Negative Provisions Effects on Various MNE Variables (Indv. Shapley)

	(1) Employee Costs	(2) Costs per employee	(3) Employees	(4) Tangible Assets	(5) Intangible Assets	(6) Revenues
			Panel I			
Positive-FDI	-0.0394***	-0.0277*	0.0120**	-0.0010	0.0111	-0.0010
Provisions (Indv. Shapley)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Negative-FDI	0.0293***	0.0287**	-0.0116**	0.0002	-0.0138*	0.0027
Provisions (Indv. Shapley)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)
			Panel II			
Positive-Trade	0.0205	-0.0187	-0.0065*	-0.0057	-0.0038	0.0009
Provisions (Indv. Shapley)	(0.02)	(0.02)	(0.00)	(0.01)	(0.01)	(0.00)
Negative-Trade	-0.0163	0.0162	0.0034	0.0031	-0.0024	0.0013
Provisions (Indv. Shapley)	(0.01)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)
Observations	18784	16187	190440	190440	190440	190440
Origin-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes	Yes	Yes
FTA/BIT controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimation uses PPML. Robust standard errors are in parentheses, clustered by country pair.
Domestic data are included.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 25: Positive and Negative Substantive Provisions, OLS and IPP-Bias-Corrected PPML

	Panel I: OLS				Panel II: IPP Correction			
	Panel I.A: Trade		Panel I.B: FDI (affiliates)		Panel II.A: Trade		Panel II.B: FDI (affiliates)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Positive-Trade	0.0086***		-0.0038***		0.0241***		-0.0019	
Provisions	(0.00)		(0.00)		(0.01)		(0.00)	
Negative-Trade	-0.0235***		0.0083***		-0.0405**		0.0054	
Provisions	(0.00)		(0.00)		(0.02)		(0.00)	
Positive-FDI		-0.0021		0.0048***		-0.0138		0.0054**
Provisions		(0.00)		(0.00)		(0.01)		(0.00)
Negative-FDI		-0.0004		-0.0033***		0.0131		-0.0020
Provisions		(0.00)		(0.00)		(0.01)		(0.00)
Observations	190440	190440	190440	190440	190440	190440	190440	190440
Pseudo R^2	0.945	0.995	0.995	0.945	0.998	0.998	0.998	0.998
Origin-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FTA/BIT controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimation uses PPML. Robust standard errors are in parentheses, clustered by country pair.
Domestic data are included. Group Shapley method is used. FDI is measured using number of affiliates.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

results are provided in the Online Supplement. There were no material differences relative to Tables 21 and 23.

Second, we re-estimated the main results for a sample restricted to 2017 (rather than 2019). This robustness test was called for because the sample from the World Bank DTA Database ended in 2017; we used their 2017 observations on the RHS for 2018 and 2019 in the main results. This estimation reports the sensitivity of our results to this issue. These alternative results are provided in the Online Supplement. Qualitatively, the results are the same. However, quantitatively the results with the shorter sample have slightly smaller coefficient estimates (in absolute values).

Third, in the Onlin Supplement we also provide comparisons of the Grouped Shapley results with the Individual Shapley results for the IPP correction addressed in Table 25. And we also provide comparisons of the Grouped Shapley results with the Individual Shapley results for the OLS estimates addressed in Table 25.

7.3 Averages per Affiliate

This section addresses quantitatively how various MNE cost, factor, and revenue variables of the *average* MNE affiliate react to DTA provisions; this is in contrast to how MNEs' affiliates in the (bilateral) aggregate respond to the variables. Instead of estimating the effects on country-level FDI variables, we first divided each variable by the number of affiliates at the country-pair level. The results, reported in Table 26, should be compared to those in Table 22.

Overall, the results in Table 26 are similar to those in Table 22. First, the coefficient estimates are identical qualitatively for all six variables for all four types of provisions across the two tables. Second, for Positive-FDI and Negative-FDI provisions, the coefficient estimates for costs per employee (the measure of $\frac{\gamma_{ilt} w^p_{lt}}{z_{lt}}$ from section 3.2) are not only of the same sign, they are not statistically different from one another; in Table 26 the coefficient estimates are smaller than those in Table 22. Third, for Positive-Trade and Negative-Trade provisions, the main difference is that the coefficient estimates for costs per employee are now *statistically significant* in Table 26 compared to those in Table 22. Fourth, a comparison of the column (3) coefficient estimates for numbers of employees shows larger and statistically significant effects of Positive-FDI and Negative-FDI provisions using averages per affiliate in Table 26 relative to those in Table 22. Fifth, in Table 22, Negative-FDI, Positive-Trade, and Negative-Trade provisions all have the expected effects on (aggregate) investment in intangible assets and are statistically significant. By contrast, in Table 26 only Positive-Trade provisions have a negative and statistically significant effect on (per affiliate) investments in intangible assets. Sixth, in both tables these four sets of provisions have no statistically significant effects on revenues nor on investments in tangible assets.

We note that the evidence suggests that the apparent mechanism in which Positive-FDI and Negative-FDI provisions influence MNEs is via *marginal cost per employee*, rather than revenues. In the Online Supplement, we show that the results using the Individual Shapley method on averages per affiliate of these FDI variables are generally similar qualitatively to those obtained using the Grouped Shapley method but differ quantitatively, cf., Online Supplement.

Table 26: Positive and Negative Provisions Effects on Various MNE Variables, Averages per Affiliate

	(1) Employee costs	(2) Cost per Employee	(3) Employees	(4) Tangible Assets	(5) Intangible Assets	(6) Revenues
			Panel I			
Positive-FDI Provisions	-0.0272*	-0.0429**	0.0391**	0.0132	-0.0055	0.0067
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Negative-FDI Provisions	0.0159*	0.0233**	-0.0311***	-0.0138	-0.0142	-0.0067
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
			Panel II			
Positive-Trade Provisions	0.0099	0.0287*	-0.0112	-0.0137	-0.0302*	-0.0087
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
Negative-Trade Provisions	-0.0214	-0.0575**	0.0130	0.0162	0.0254	0.0124
	(0.03)	(0.03)	(0.01)	(0.01)	(0.03)	(0.01)
Observations	42288	18134	42288	42288	42288	42288
Origin-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes	Yes	Yes
FTA/BIT controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimation uses PPML. Robust standard errors are in parentheses, clustered by country pair. Domestic data are included.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7.4 Further Decomposition of Positive and Negative Trade and FDI Provisions: Multilateral vs. Bilateral Provisions

The authors examined all 164 substantive provisions to interpret the provisions as either bilateral (affecting only signing parties) or multilateral (affecting third countries) in "nature."[36] We note several illustrative provisions. For instance, consider first provisions for Export Taxes. Export Tax Provision 03 asks whether new export taxes or quantitative restrictions are prohibited "between the Parties" (referring to the particular pair of countries);

[36] See the Online Supplement for further details.

a 1 indicates the bilateral prohibition of such new taxes or restrictions (and hence is a liberalization) between the pair. By contrast, Export Tax Provision 04 asks whether the exporter prohibits voluntary export constraints “inconsistent with GATT Article VI”; hence, this provision has a multilateral (liberalization) nature since it would apply to all potential exporters.

Consider also representative Investment Provisions. Investment Provision 28 asks whether the origin country’s Fair and Equitable Treatment (FET) clause provides that the finding of an FET violation take into account the “level of development of the host country”; this provision has a bilateral (liberalization) nature since it is home and host country specific. By contrast, Investment Provision 22 asks whether the agreement “grants Fair and Equitable Treatment (FET)”; this provision has a multilateral nature to it.

We use our classification of bilateral and multilateral provisions described in the Online Supplement to estimate their effects on trade and FDI. Specifically, we split the Shapley Positive-Trade, Negative-Trade, Positive-FDI and Negative-FDI groups of provisions into their multilateral and bilateral classifications.[37]

We report the results using the Grouped Shapley approach in Table 27; the Online Supplement reports the results using the Individual Shapley approach (which are largely similar). Regarding Positive-Trade provisions in Table 27, both bilateral ones and multilateral ones have the expected positive effects on trade; note that bilateral provisions have a quantitatively larger partial effect. Bilateral Negative-Trade provisions have a significant negative effect on trade, as expected; however, multilateral Negative-Trade provisions do not have a significant effect on trade.

Similar to earlier, bilateral Positive-FDI (Negative-FDI) provisions have significantly negative and positive effects, respectively, on trade. Interestingly, multilateral Positive-FDI provisions have a positive effect on trade, suggesting that trade and FDI tend to be complements to each other with respect to multilateral Positive-FDI provisions. By contrast, multilateral Negative-FDI provisions have the same positive effect on trade as bilateral Negative-FDI provisions.

Of the Positive- and Negative-Trade provisions, only multilateral Negative-Trade provisions have a statistically significant effect on FDI; note that the effect is as expected. In contrast, only bilateral Positive-FDI provisions have the expected positive and statistically significant effects on FDI.

In general, we find that the Positive-Trade “bilateral” provisions have quantitatively larger effects on trade than “multilateral” ones. Positive-FDI bilateral provisions have a strong negative impact on trade, but Positive-FDI multilateral provisions have a significant positive effect on trade. Only Negative-Trade multilateral provisions have significant positive

[37] Recall that the latter classifications are based on our reading of each provision.

effects on FDI. By contrast, only Positive-FDI bilateral provisions have significant positive effects on FDI.

Table 27: Robustness: Multilateral vs Bilateral provisions

	Panel A: Trade		Panel B: FDI (affiliates)	
	(1)	(2)	(3)	(4)
Positive-Trade provisions (Bilateral)	0.0348***		0.0001	
	(0.01)		(0.00)	
Negative-Trade provisions (Bilateral)	-0.0274***		0.0051	
	(0.01)		(0.00)	
Positive-Trade provisions (Multilateral)	0.0100**		-0.0023	
	(0.00)		(0.00)	
Negative-Trade provisions (Multilateral)	0.0036		0.0073**	
	(0.01)		(0.00)	
Positive-FDI provisions (Bilateral)		-0.0614***		0.0099**
		(0.02)		(0.00)
Negative-FDI provisions (Bilateral)		0.0094**		-0.0014
		(0.00)		(0.00)
Positive-FDI provisions (Multilateral)		0.0165***		0.0021
		(0.00)		(0.00)
Negative-FDI provisions (Multilateral)		0.0181*		-0.0029
		(0.01)		(0.00)
Observations	190440	190440	190440	190440
Pseudo R^2	0.998	0.998	0.998	0.998
Origin-Year FE	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes
FTA/BIT controls	Yes	Yes	Yes	Yes

Notes: Estimation uses PPML. Robust standard errors are in parentheses, clustered by country pair. Domestic data are included. Group Shapley method is used.
* p<0.10, ** p<0.05, *** p<0.01

7.5 Spillover Effects of Provisions

An issue of long interest in research on FDI and MNEs' behavior is the possible role of *spillover* effects. In our context, spillovers refer to the effects of changes in certain factors on third-country FDI. As just one representative analysis, Gil-Pareja et al. (2022) examined empirically among other issues the potential spillover effects of the presence of regional headquarters in some country j of foreign MNEs (based, say, in i) on performance of domestic firms in j as well as on other (third-country) firms operating in j.

Up to now, we have examined the partial effects of provisions between countries i and j on bilateral trade and FDI between i and j. In this section, we examine the effects of provisions in some destination country j – and, in particular, *excluding* bilateral provisions between i and j – on trade and FDI between i and j. Such an approach was used in Gil-Pareja et al. (2022) to estimate the effects of regional headquarters of MNEs based abroad in some country i on domestic firms and on affiliate firms of non-i foreign MNEs.

In the context of our study, we construct a variable that is the sum of provisions signed by destination country j *minus* the bilateral provisions signed between countries i and j. We do this for positive-trade and negative-trade provisions as well as positive-FDI and negative-FDI provisions. Hence, for positive-trade provisions, we refer to the new variable as "Positive-Trade Provisions (Sum at Destination)"; analogous names are used for the other three category types. By including Positive-Trade Provisions (Bilateral) alongside Positive-Trade Provisions (Sum at Destination), we can decompose the partial effects of Positive-Trade Provisions (on trade or FDI) between their direct bilateral effects and their third-country spillover effects.[38]

The results of this alternative specification using the Grouped Shapley approach are shown in Table 28. For trade, shown in Panel A, one more Positive-Trade Provision (Sum at Destination) has approximately one-tenth the positive impact on trade as a Positive-Trade (bilateral) Provision. Similarly, Negative-Trade spillover effects are smaller than Negative-Trade bilateral effects (in absolute value), as seen in column (1). Looking to column (2), spillover effects on bilateral trade from FDI Sum-at-Destination provisions also tend to be much smaller than partial effects of bilateral FDI provisions. Interestingly, all eight variables in Panel A have the expected signs and the coefficient estimates are statistically significant.

Panel B reports the results for the effects on bilateral FDI of bilateral provisions and spillover effects. Interestingly, in column (3) we find that the only statistically significant positive and negative trade provisions affecting bilateral FDI are the Sum-at-Destination variables. Analogously, the Sum-at-Destination Positive-FDI and Negative-FDI variables have the expected effects and are statistically significant, as shown in column (4). Hence, Sum-at-Destination Positive-FDI Provisions have positive "spillover" effects on bilateral FDI; Sum-at-Destination Negative-FDI Provisions have negative "spillover" effects on bilateral FDI. These results are consistent with other evidence that spillover effects are non-trivial for FDI.

[38] Following Heid et al. (2021), to avoid multicollinearity with the fixed effects, the variable Provisions (Sum at Destination) is interacted with the border dummy.

Table 28: Robustness: Bilateral vs. Sum-at-Destination Provisions

	Panel A: Trade		Panel B: FDI (affiliates)	
	(1)	(2)	(3)	(4)
Positive-Trade Provisions	0.0179***		0.0002	
	(0.00)		(0.00)	
Negative-Trade Provisions	-0.0339***		0.0001	
	(0.01)		(0.00)	
Positive-Trade Provisions (Sum at destination)	0.0022***		-0.0004***	
	(0.00)		(0.00)	
Negative-Trade Provisions (Sum at destination)	-0.0043***		0.0011***	
	(0.00)		(0.00)	
Positive-FDI Provisions		-0.0257***		0.0024
		(0.01)		(0.00)
Negative-FDI Provisions		0.0142**		-0.0009
		(0.01)		(0.00)
Positive-FDI Provisions (Sum at destination)		-0.0009**		0.0003**
		(0.00)		(0.00)
Negative-FDI Provisions (Sum at destination)		0.0011***		-0.0003***
		(0.00)		(0.00)
Observations	190440	190440	190440	190440
Pseudo R^2	0.987	0.987	0.961	0.961
Origin-Year FE	Yes	Yes	Yes	Yes
Destination-Year FE	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes
FTA/BIT controls	Yes	Yes	Yes	Yes

Notes: Estimation uses PPML. Robust standard errors are in parentheses, clustered by country pair. Domestic data are included. Group Shapley method is used.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In the Online Supplement, we provide the analogous results using the Individual Shapley method. Those results are qualitatively the same and quantitatively similar.

8 General Equilibrium Comparative Statics

As one might expect, there are numerous potential numerical general equilibrium comparative static exercises one can perform using our novel MREID data set and the DTA provisions, along with their estimated partial effects on trade and/or various FDI variables. However, due to the length of this paper, we limit ourselves to two representative counterfactual exercises.[39]

Since most counterfactual exercises with regard to trade agreements focus on trade effects, both exercises will focus on the general equilibrium export, nominal GDP, consumer prices, and producer prices effects of Positive-Trade and Negative-Trade provisions. In Counterfactual 1, we calculate the general equilibrium effects of *removing* all Positive-Trade Provisions (using the Grouped Shapley estimates). In Counterfactual 2, we calculate the general equilibrium effects of removing all Negative-Trade Provisions.

To concisely generate the two representative counterfactuals, we adopt the methodology based upon the General Equilibrium Poisson Pseudo Maximum Likelihood (GEPPML) framework explained in Anderson et al. (2018). Due to our paper's length, we summarize briefly the methodology, referring the reader to Anderson et al. (2018), Yotov et al. (2016) and Anderson et al. (2020) for details. This framework provides a well established methodology for incorporating the partial equilibrium estimates of trade effects of trade policies into a standard international trade general equilibrium framework suitable for numerical counterfactuals. In the context of trade, the essence of the approach is to first use a structural gravity equation – such as equation (12) in section 6.1 – to estimate a change in $t_{ijt}^{1-\sigma}$, where σ is the elasticity of substitution in consumption.[40] Equation (12) implies that the ratio of predicted bilateral trade to its benchmark friction-less flow is equal to a power tranform of the ratio of bilateral trade costs to the product of the outward and inward multilateral price (resistance) terms. Theoretically, the multilateral prices are obtained as the solution to a nonlinear pair of equations derived from global market clearances and each country's budget constraint being met. Fally (2015) showed that – when equation (12) is estimated using PPML (as we do in our paper) – the estimated exporter and importer fixed effects are *precisely equal* to the multilateral prices that satisfy the general equilibrium system of equations.

[39] In our GE experiments we do not include sectoral dimensions because we want to showcase the effect of the estimation biases. In this case, introducing sectoral heterogeneity is not particularly critical (Giri et al., 2021).

[40] There are many theoretical approaches for grounding the trade elasticity; the approach we use here is flexible to all of them.

In Counterfactual 1, the counterfactual is the removal of *all Positive-Trade Provisions*. As we know from our econometric estimates, Positive-Trade Provisions have an estimated positive *partial* effect on bilateral trade. Using the GEPPML Structural Gravity methodology, the top four figures in Figure 7 report for the 138 countries in our (ITPD-E) trade data set the percentage changes for each country for four variables: exports, (nominal) GDPs, consumer price levels, and producer price levels.

The removal of all Positive-Trade Provisions – leaving present the (existing) Negative-Trade Provisions – increases trade for some countries, but the vast majority of countries' exports decline as shown in Figure 7. Calculation of the change in the world's export shows a decline of 4.80%. The elimination of Positive-Trade Provisions reduces the world's nominal GDP by 0.77%, with only 25 of the 138 countries showing higher GDP. Since world consumer prices rise by 0.65%, world output (real income) declines by 1.42%. We also can infer based on the second row of figures, alongside the first row, that *real incomes* of households in the overwhelming majority of countries decline as consumer prices are higher in almost all countries; only three countries see declines in consumer prices. By contrast, we note the improvement of producers' prices in a majority of countries.

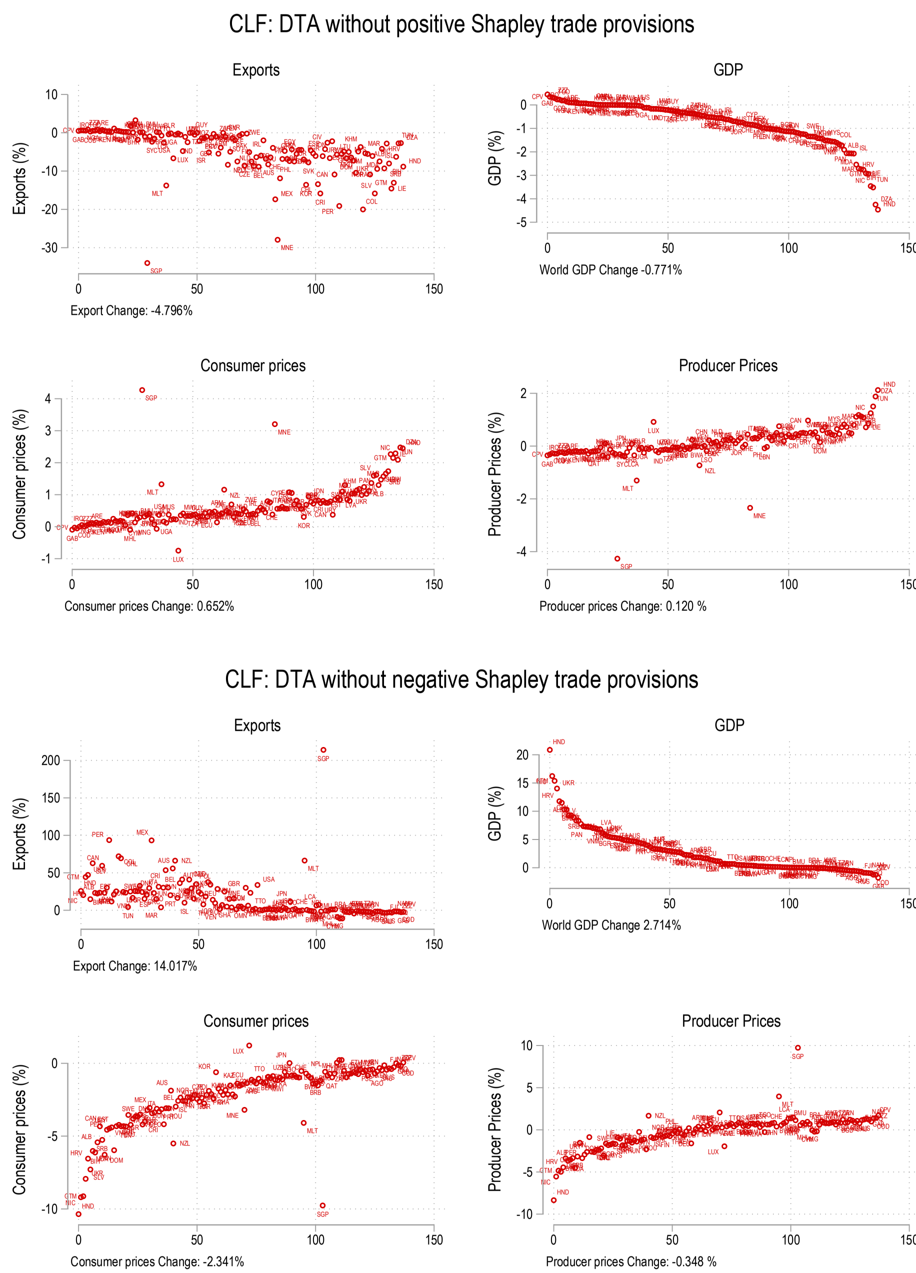
The detailed percentage changes for the world and for all 138 countries for Counterfactual 1 is provided in the Online Supplement Appendix. For instance, for the United States (USA), real income declines by 0.49 of 1 percent as U.S. exports decline by 6.67 percent. Moreover, U.S. firms incur a loss in producer prices of 0.25 of 1 percent.

In the second counterfactual reported in the lower panel of Figure 7, we compute the effects of including only Positive-Trade Provisions in DTAs. That is, we *remove* all *Negative-Trade* Provisions. In this case, the world's output *increases* by 5.05%, as world nominal GDP increases by 2.71% and consumer prices decline by 2.34%. Furthermore, producers' prices worldwide improve by 0.35% as world exports surge by 14.02%.

As documented in Online Supplement, the vast majority of countries' exports and nominal GDPs increase and consumer price levels decline. For instance, removing all Negative-Trade Provisions increases U.S. real income by *more than 2 percent* (2.06%) following increases in U.S. exports of *33.76%*, with U.S. nominal GDP increasing 0.67% alongside a decline in U.S. consumer prices of 1.38%.

As we know from our earlier analysis, there is considerable heterogeneity in individual provisions' partial effects on world trade. Our decomposition of provisions' effects into positive and negative effects is just one categorization of many possible ones that highlight this heterogeneity. In our counterfactuals, Negative-Trade Provisions having approximately *three times* the (diminishing) effect on world output as Positive-Trade Provisions have on (augmenting) world output.

Figure 7: Counterfactuals



Notes: Counterfactuals using GEPPML

9 Conclusions

Several years ago, the World Bank released its "Deep Trade Agreements" database, which included a large panel of binary indicators of nearly 1,000 provisions in the economic integration agreements among nearly all the countries in the world. Beyond categorizing provisions as substantive vs. non-substantive, the World Bank decomposed substantive provisions as either *liberalizations* or *obligations*. While categorization of a provision as a liberalization suggests it would likely increase trade between countries, provisions categorized as obligations suggests – but does not necessitate – that these provisions might impede trade, even though such provisions are included to address non-trade goals such as to guarantee minimum labor standards or reduce damage to the environment.

In this paper, we used a novel classification method based upon the Shapley Value technique from cooperative game theory to provide *ex post* quantitative estimates of these provisions' impacts, with a key finding that many of the World Bank's provisions categorized as liberalizations increase trade, but many decrease trade. Similarly, many of the World Bank's provisions categorized as obligations decrease trade, but many increase trade. Furthermore, using a new data set (MREID) from the U.S. International Trade Commission on MNE activities (which the authors contributed to), we find similar *ex post* effects for FDI.

Among numerous empirical results, we first find evidence that DTA provisions positively affect trade (FDI), with one additional randomly selected provision in the World Bank's data set (with nearly 1,000 provisions) increasing *aggregate* trade (FDI) by 0.06 (0.01) of 1 percent. Second, the marginal effect of a "substantive" provision on trade (FDI) is four (three) times that of a non-substantive provision. Third, DTAs with the mean number of substantive provisions (30) increase trade by 11 percent and increase FDI by 2 percent. Fourth, substantive provisions have quadratic effects on trade and FDI, but the nature of the quadratic relationship differs between them. Fifth, when all individual provisions that have positive (negative) effects on FDI are grouped, these sets of provisions positively (negatively) affect FDI *but* negatively (positively) affect trade. Furthermore, when all individual provisions that have positive (negative) effects on trade are grouped, these sets of provisions positively (negatively) affect trade *but* negatively (positively) affect FDI. Sixth, we find unbiased estimates of the signs of individual provisions' effects on trade and FDI, noting that the World Bank's "liberalization provisions" can have positive or negative effects and their "obligation provisions" can have negative or positive effects. Seventh, owing to the breadth of data in our MREID database, we identify channels through which such effects work; for example, we find statistically significant evidence that sets of provisions that positively (negatively) affect FDI are decreasing (increasing) marginal *costs per employee*.

In an extensive robustness analysis, we find that our results are qualitatively the same whether we use the Individual or Grouped Shapley Value approach, PPML or OLS, and bilateral aggregate data or averages per affiliate. Furthermore, we find evidence that Positive-FDI provisions work largely through multilateral rather than bilateral channels and that positive FDI provisions have (third-country) spillover effects.

Finally, using two numerical general equilibrium comparative static exercises, we find a large quantitative difference in the economic effects of provisions that increase trade (arguably, "true" liberalizations) and those that decrease trade (arguably, "true" obligations). In our example, the negative effects on world output of negative-trade provisions are three times the size of the positive effects on world output of positive-trade provisions. The methodology in this paper provides a framework for analyzing numerous other categorizations of individual provisions to better understand qualitatively and quantitatively the economic impacts on trade and measures of MNEs' activities (including FDI in affiliates) of numerous non-tariff provisions at the core of the world's deep "trade" agreements.

References

Ahmad, S., J. Bergstrand, J. Paniagua, and H. Wickramarachi (2025). The multinational revenue, employment, and investment database (MREID). *Review of International Economics*.

Anderson, J., M. Larch, and Y. Yotov (2019). Trade and investment in the global economy: A multi-country dynamic analysis. *European Economic Review 120*.

Anderson, J., M. Larch, and Y. Yotov (2020). Transitional growth and trade with frictions: A structural estimation framework. *The Economic Journal 130*, 1583–1607.

Anderson, J. and E. van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review 93*(1), 170–192.

Anderson, J. E., M. Larch, and Y. V. Yotov (2018). Geppml: General equilibrium analysis with ppml. *The World Economy 41*(10), 2750–2782.

Antràs, P., E. Fadeev, T. C. Fort, and F. Tintelnot (2024). Export-platform FDI: Cannibalization or complementarity? In *AEA Papers and Proceedings*, Volume 114, pp. 130–135. American Economic Association.

Antràs, P. and E. Helpman (2008). *Contractual Frictions and Global Sourcing,* in: Helpman, E., D. Marin, and T. Verdier (eds.), The Organization of Firms in a Global Economy. Harvard University Press.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012). New trade models, same old gains? *American Economic Review 102*(1), 94–130.

Arkolakis, C., N. Ramondo, A. Rodriguez-Clare, and S. Yeaple (2018). Innovation and production in the global economy. *American Economic Review 108*(8), 2128–2173.

Baier, S. and J. Bergstrand (2007). Do free trade agreements actually increase members' international trade? *Journal of International Economics 71*(1), 72–95.

Baier, S., J. Bergstrand, and M. Clance (2018). Heterogeneous effects of economic integration agreements. *Journal of Development Economics 135*, 587–608.

Baier, S., J. Bergstrand, and M. Feng (2014). Economic integration agreements and the margins of international trade. *Journal of International Economics 93*(2), 339–350.

Baier, S., A. Kerr, and Y. Yotov (2017). Gravity, distance, and international trade. In B. Blonigen and W. Wesley (Eds.), *Handbook of International Trade and Transportation*. Edward Elgar Publishing.

Baier, S. and N. Regmi (2023). Using machine learning methods to capture heterogeneity in trade agreements. *Open Economies Review 34*, 863–894.

Bergstrand, J. and P. Egger (2007). A knowledge-and-physical-capital model of international trade flows, foreign direct investment, and multinational enterprises. *Journal of International Economics 73*, 278–308.

Bergstrand, J. and P. Egger (2010). A general equilibrium theory for estimating gravity equations of bilateral fdi, final goods trade, and intermediate goods trade. In S. Brakman and P. van Bergeijk (Eds.), *The Gravity Model in International Trade: Advances and Applications*, pp. 29–70. Cambridge, UK: Cambridge University Press.

Bergstrand, J. and P. Egger (2013). Shouldn't physical capital *also* matter for multinational enterprise activity? *Review of International Economics 21*, 945–965.

Bergstrand, J. and P. Egger (2014). What determines bits? *Journal of International Economics 90*(1), 107–122.

Bergstrand, J., M. Larch, and Y. Yotov (2015). Economic integration agreements, border effects, and distance elasticities in the gravity equation. *European Economic Review 78*, 307–327.

Borchert, I., M. Larch, S. Shikher, and Y. V. Yotov (2021). The international trade and production database for estimation (itpd-e). *International Economics 166*, 140–166.

Borchert, I., M. Larch, S. Shikher, and Y. V. Yotov (2022a). Disaggregated gravity: Benchmark estimates and stylized facts from a new database. *Review of International Economics 30*(1), 113–136.

Borchert, I., M. Larch, S. Shikher, and Y. V. Yotov (2022b). The international trade and production database for estimation-release 2. *USITC Working Paper 2022–07–A*.

Breinlich, H., V. Corradi, N. Rocha, M. Ruta, J. Santos Silva, and T. Zylkin (2021). Using machine learning to assess the impact of deep trade agreements. In A. Fernandes, N. Rocha, and M. Ruta (Eds.), *The Economics of Deep Trade Agreements*, pp. 25–36. London, UK: Centre for Economic Policy Research.

Burchardi, K. B., T. Chaney, and T. A. Hassan (2019). Migrants, ancestors, and foreign investments. *The Review of Economic Studies 86*(4), 1448–1486.

Buthe, T. and H. Milner (2014). Foreign direct investment and institutional diversity in trade agreements: Credibility, commitment, and economic flows in the developing world. *Social Science Electronic Publishing 6*(4), 88–122.

Casella, B., M. Borga, and M. K. Wacker (2023). Measuring multinational production with foreign direct investment statistics: Recent trends, challenges, and developments. *IMF Working Papers* (2023/113).

Choi, J., J. Hyun, and Z. Park (2024). Bound by ancestors: Immigration, credit frictions, and global supply chain formation. *Journal of International Economics 147*, 103855.

Cuadros, A., J. Martín-Montaner, and J. Paniagua (2019). Migration and FDI: The role of job skills. *International Review of Economics & Finance 59*, 318–332.

Darlington, R. (1968). Multipleregression in psychological research and practice. *Psychological Bulletin 79*, 161–182.

Dhingra, S., R. Freeman, and E. Mavroeidi (2018“). Beyond tariff reductions: What extra boost to trade from agreement provisions? *London School of Economics Centre for Economic Performance Discussion Paper*.

Egger, P., A. Pirotte, and C. Titi (2023). International investment agreements and foreign direct investment: A survey. *The World Economy*.

Fally, T. (2015). Structural gravity and fixed effects. *Journal of International Economics 97*, 76–85.

Fontagne, L., N. Rocha, M. Ruta, and G. Santoni (2022). The economic impact of deep trade agreements. *CESifo Working Paper* (9529).

Garcia-Bernardo, J., J. Fichtner, F. W. Takes, and E. M. Heemskerk (2017). Uncovering offshore financial centers: Conduits and sinks in the global corporate ownership network. *Nature (Scientific Reports) 7*(1), 1–10.

Gil-Pareja, S., R. Llorca-Vivero, and J. Paniagua (2022). Regional headquarters and foreign direct investment. *Review of International Economics 30*(4), 1288–1310.

Giri, R., K.-M. Yi, and H. Yilmazkuday (2021). Gains from trade: Does sectoral heterogeneity matter? *Journal of International Economics 129*, 103429.

Glennon, B. (2023). How do restrictions on high-skilled immigration affect offshoring? evidence from the h-1b program. *Management Science*.

Gopinath, G., Ş. Kalemli-Özcan, L. Karabarbounis, and C. Villegas-Sanchez (2017). Capital allocation and productivity in south europe. *The Quarterly Journal of Economics 132*(4), 1915–1967.

Gould, D. M. (1994). Immigrant links to the home country: empirical implications for us bilateral trade flows. *The Review of Economics and Statistics*, 302–316.

Grounder, A., R. Falvey, and G. Rajaguru (2019). The effects of preferential trade agreements on foreign direct investment: Evidence from the african-caribbean-pacific region. *Open Economies Review 30*.

Guvenen, F., R. M. Jr., D. Rassier, and K. Ruhl (2022). Offshore profit shifting and aggregate measurement: Balance of payment, foreign investment, productivity, and the labor share. *American Economic Review 112*(6), 1848–1884.

Harris, R. (1975). *A Primer of Multivariate Statistics*. New York: Academic Press.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), *Handbook of Interantional Economics, Volume 4*. Amsterdam: Elsevier.

Heid, B., M. Larch, and Y. V. Yotov (2021). Estimating the effects of non-discriminatory trade policies within structural gravity models. *Canadian Journal of Economics/Revue canadienne d'économique 54*(1), 376–409.

Herman, P. (2024). In too deep? the economic impacts of deep trade agreements. *USITC Office of Economics Working Paper 2024–12–D*.

Hofmann, C., A. Osnago, and M. Ruta (2017). Horizontal depth: A new database on the content of preferential trade agreements. *World Bank Working Paper 7981*.

Horn, H., P. Mavroidis, and A. Sapir (2010). Beyond the wto? an anatomy of eu and us preferential trade agreements. *The World Economy 33*(11), 1565–1588.

Kalemli-Ozcan, S., B. Sorensen, C. Villegas-Sanchez, V. Volosovych, and S. Yesiltas (2015). How to construct nationally representative firm-level data from the orbis global database: New facts and aggregate implications. Technical report, National Bureau of Economic Research.

Kalemli-Özcan, S., B. E. Sørensen, C. Villegas-Sanchez, V. Volosovych, and S. Yesiltas (2023). How to construct nationally representative firm level data from the orbis global database: New facts on smes and aggregate implications for industry concentration. *American Economic Journal: Macroeconomics*.

Kohl, T., S. Brakman, and H. Garretsen (2016). Do trade agreements stimulate international trade differently? evidence from 296 trade agreements. *The World Economy 39*(1), 97–131.

Kox, H. L. and H. Rojas-Romagosa (2019). Gravity estimations with fdi bilateral data: Potential fdi effects of deep preferential trade agreements. *Robert Schuman Centre for Advanced Studies Research Paper No. RSCAS 2019/70 70*.

Laget, E., N. Roch, and G. Varela (2021). Deep trade agreement and foreign direct investments. *World Bank Policy Research Working Paper 9829*.

Larch, M. and Y. V. Yotov (2025). Deep trade agreements and FDI in partial and general equilibrium: a structural estimation framework. *The World Bank Economic Review 39*(2), 281–307.

Limao, N. (2016). Preferential trade agreements. In K. Bagwell and R. Staiger (Eds.), *Handbook of Commerical Policy*. Elsevier Science.

Lipovetsky, S. and M. Conklin (2001). Analysis of regression in game theory approach. *Applied Stochastic Models in Business and Industry 17*, 319–330.

Markusen, J. (2002). *Multinational Firms and the Theory of International Trade*. Cambridge, Massachusetts: The MIT Press.

Mattoo, A., A. Mulabdic, and M. Ruta (2017). Trade creation and trade diversion in deep agreements. *World Bank Policy Research Working Paper*.

Mattoo, A., N. Rocha, and M. Ruta (2020). *Handbook of Deep Trade Agreements*. World Bank, Washington, DC.

Mistura, F. and C. Roulet (2019). The determinants of foreign direct investment: Do statutory restrictins matter? *OECD Working Papers on International Investment*.

Morales, N. (2023). High-skill migration, multinational companies and the location of economic activity. *Review of Economics and Statistics*, 1–45.

Mulabdic, A., A. Osnago, and M. Ruta (2017). Deep integration and uk-eu trade relations. *World Bank Policy Research Working Paper*.

Olden, A. and J. Møen (2022). The triple difference estimator. *The Econometrics Journal 25*(3), 531–553.

Ornelas, E. and J. L. Turner (2024). The costs and benefits of rules of origin in modern free trade agreements. *Journal of International Economics*, 103874.

Osnago, A., N. Rocha, and M. Ruta (2019). Deep trade agreements and vertical fdi: The devil is in the details. *Canadian Journal of Economics/Revue canadienne d'économique 52*(4), 1558–1599.

Ottaviano, G. I., G. Peri, and G. C. Wright (2018). Immigration, trade and productivity in services: Evidence from uk firms. *Journal of International Economics 112*, 88–108.

Paniagua, J., E. Figueiredo, and J. Sapena (2015). Quantile regression for the FDI gravity equation. *Journal of Business Research 68*(7), 1512–1518.

Peri, G. and F. Requena-Silvente (2010). The trade creation effect of immigrants: evidence from the remarkable case of spain. *Canadian Journal of Economics/Revue canadienne d'économique 43*(4), 1433–1459.

Ramondo, N. and A. Rodriguez-Clare (2013). Trade, multinational production, and the gains from openness. *Journal of Political Economy 121*(2), 273–322.

Santacreu, A. M. (2021). Dynamic gains from trade agreements with intellectual property provisions. *FRB St. Louis Working Paper* (2021-10).

Santos Silva, J. and S. Tenreyro (2006). The log of gravity. *Review of Economics and Statistics 88*(4), 641–658.

United States International Trade Commission (2016). *Economic Impact of Trade Agreements Implemented under Trade Authorities Procedures, 2016 Report*. Washington, DC: U.S. International Trade Commission.

United States International Trade Commission (2021). *Economic Impact of Trade Agreements Implemented under Trade Authorities Procedures, 2021 Report*. Washington, DC: U.S. International Trade Commission.

Vehtari, A., A. Gelman, and J. Gabry (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and computing 27*, 1413–1432.

Weidner, M. and T. Zylkin (2021). Bias and consistency in three-way gravity models. *Journal of International Economics 132*.

Yotov, Y., R. Piermartini, J. Monteiro, and M. Larch (2016). *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*. Geneva, Switzerland: The World Trade Organization.