

CHAPTER 6

Preferential Trade Agreements

N. Limão

University of Maryland, College Park, MD, United States

Contents

1. Introduction	281
2. Stylized Facts and a Taxonomy	284
2.1 A Definition	284
2.2 Common Classification of PTAs and Their Evolution Over Time	285
2.3 A Taxonomy of “Modern” PTAs	287
2.3.1 <i>Depth in Economic Policy Cooperation</i>	288
2.3.2 <i>Breadth in Economic Policy Cooperation</i>	290
2.3.3 <i>Complementarity of Depth and Economic Breadth</i>	291
2.3.4 <i>Breadth in Noneconomic Policy Cooperation</i>	292
2.4 Trends in Modern PTAs	292
2.4.1 <i>Emerging Dimensions and Complementary Data</i>	293
2.5 Stylized Facts	293
3. Trade and Welfare Effects on Members	294
3.1 Ex Post Trade Effects: Naïve Gravity and a PTA Formation “Puzzle”	295
3.1.1 <i>Gravity Approach</i>	295
3.1.2 <i>Average Partial Effects and Selection Bias</i>	296
3.1.3 <i>Naïve Gravity</i>	297
3.2 Ex Post Trade Effects: Theory Consistent Estimates	300
3.2.1 <i>Structural Gravity and Multilateral Trade Determinants</i>	300
3.2.2 <i>Endogenous PTAs and Bilateral Trade Determinants</i>	302
3.2.3 <i>Sample Issues and Heterogeneous PTA Effects</i>	305
3.2.4 <i>Summary</i>	306
3.3 A PTA Trade Elasticity Puzzle?	307
3.3.1 <i>The Magnitude and Timing of Aggregate PTA Effects</i>	307
3.3.2 <i>Puzzle: Definition and Existing Evidence</i>	308
3.3.3 <i>Puzzle: Novel Evidence</i>	309
3.3.4 <i>Features of Possible Explanations for the Elasticity Puzzle</i>	312
3.4 Ex Post Estimates of WTO Trade Effects	313
3.5 General Equilibrium Trade and Welfare Effects of PTAs	315
4. Economic Effects of Deeper PTAs	317
4.1 Nontariff Barriers	317
4.2 Intermediates and Vertical Integration	319
4.3 Trade Policy Uncertainty (TPU) as a Motive for Trade Agreements	321
4.3.1 <i>Sources of TPU</i>	322
4.3.2 <i>Direct Evidence of TPU Reductions via Agreements</i>	324
4.4 A TPU-Investment Mechanism	325

4.4.1	<i>Mechanism</i>	325
4.4.2	<i>TPU Augmented Industry Gravity</i>	326
4.5	Ex Post Trade and Firm/Variety Entry Investment Estimates of Deeper PTAs	328
4.5.1	<i>Expanding and Securing Existing Preferences</i>	328
4.5.2	<i>Average Treatment vs Policy Effects</i>	329
4.5.3	<i>PTAs as Insurance Against Trade Wars</i>	329
4.6	Other Evidence and Future Work	331
5.	PTA Formation and Policies: Motives and Determinants	332
5.1	Motives and Mechanisms	332
5.1.1	<i>Traditional</i>	332
5.1.2	<i>Nontraditional</i>	337
5.2	PTA and Preference Determinants	342
5.2.1	<i>Economic Determinants of PTAs Under Exogenous Tariffs: Binary Choice Approach</i>	342
5.2.2	<i>Alternative Determinants and Approaches to PTA Formation Under Endogenous Tariffs</i>	345
5.2.3	<i>Summary</i>	347
6.	Agreement and Policy Interdependence	348
6.1	Agreement Interdependence: Approaches and Open Questions	348
6.2	Policy Interdependence	350
6.2.1	<i>Multilateral Tariff Interdependence</i>	351
6.2.2	<i>Additional Dimensions of Interdependence</i>	355
6.3	Wither Multilateral Tariffs, Wither Diversion, and Nonmember Discrimination?	356
7.	Lessons and Future Research	357
7.1	Deeper and Broader Policy Cooperation	357
7.2	Trade-Related Effects on Members	358
7.3	PTA Formation and Policies	358
7.4	Interdependence	359
	Acknowledgments	360
	References	360

Abstract

A large and growing number of countries participate in multiple preferential trade agreements (PTAs), which increasingly entail broad cooperation over policies extending far beyond trade barriers. I review the traditional and nontraditional motives for PTAs and their empirical determinants as well as their impacts on trade and on multilateral liberalization. I argue that the broad nature of modern PTAs, their substantial creation of bilateral trade, and their modest effects on members' tariffs, require us to augment the economic and policy structure of traditional models of PTAs as a static preferential tariff reduction. Throughout I draw lessons from the existing literature and point toward many interesting paths for future research, to advance our understanding of the causes of modern PTAs and their impacts on trade-related outcomes and beyond.

Keywords

Trade, Preferential agreements, WTO, Globalization

JEL Classification Codes

F02, F1, F4, F5, F6

1. INTRODUCTION

In 2010 the number of preferential trade agreements (PTAs) in force was four times higher than in 1990. The participation in PTAs is widespread: in 2010 each member of the World Trade Organization (WTO) also participated in an average of 13 PTAs, up from only 2 in 1990 (WTO, 2011). This trend, the negotiation of mega-agreements by the United States and Europe and the evidence discussed later, indicates that PTAs are the most important source of trade policy reform in the last 20 years for most countries.

In Fig. 1, we see that the proliferation of PTAs has continued after the creation of the WTO in a period when nonpreferential MFN tariffs were declining. Some of the largest growth has occurred in the last decade even though average MFN tariffs are at their lowest, averaging less than 8% in 2009. The traditional Vinerian view of PTAs, and most of the economic analysis, treats them as a static reduction in tariffs with respect to a preferential partner. But if the initial tariffs are already low then so is the preferential tariff margin, which raises two basic questions. What explains the formation and proliferation of so many PTAs and what are their basic trade and welfare effects on members?

To answer these two questions, I first provide some stylized facts about the importance and evolution of trade between PTA members. Their share of world trade almost tripled between 1965 and 2010, with “deeper” PTAs becoming increasingly more important. A detailed examination of the provisions of modern PTAs in 2011 reveals policy cooperation far beyond reductions in applied tariffs. I provide a taxonomy of PTAs in terms of policy depth and breadth, where the latter includes economic and noneconomic

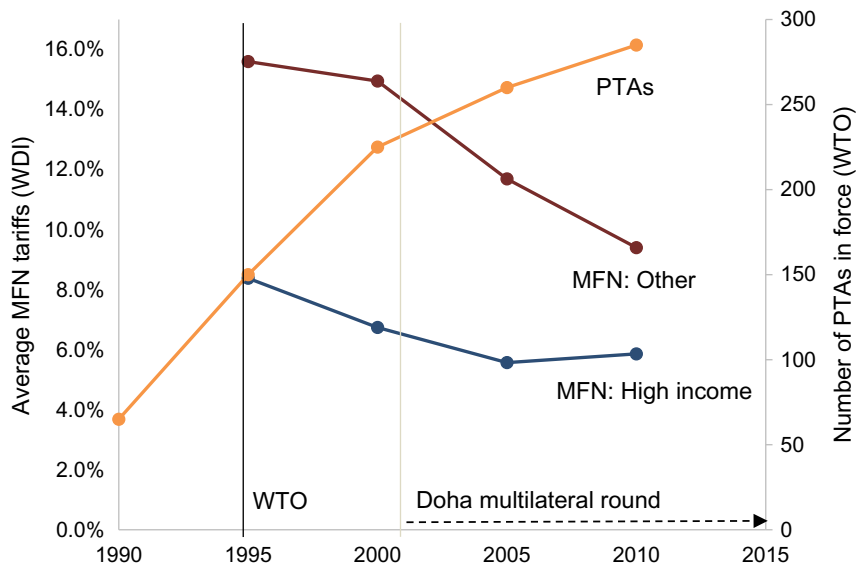


Fig. 1 Preferential and multilateral liberalization.

provisions. Some of these provisions also evolved over time in the context of the GATT/WTO and others go far beyond it.

Despite the diverse nature of these agreements, they share one common feature, a policy that aims to increase market access for at least one member. Therefore in [Section 3](#), I examine if PTAs cause increases in bilateral trade between members. After discussing the methodological issues associated with these estimates I conclude that, when properly estimated, these effects are large on average; possibly too large to be explained by the observed preferential tariff reductions on final goods. Moreover, the effects are heterogeneous across PTAs, even after controlling for tariffs, and increasing up to 10 years after the agreement, suggesting a gradual or dynamic effect. From the perspective of the traditional view of PTAs as static tariff reductions these facts appear puzzling since the observed tariff reductions are modest, as the evidence shows for PTAs since 1990. I describe what features of a richer economic and policy setting would explain the “puzzle.”

In [Section 4](#), I examine specific economic motives and effects of deeper PTAs, which address trade policies beyond tariffs and aim to integrate production structures across countries. These features of recent deeper PTAs augment the economic and policy structure relative to the traditional view in a way that can help to explain the estimated aggregate trade effects. I argue the trade policy structure should be augmented to incorporate current nontariff barriers (NTBs) and also uncertainty about future policies, where the latter is particularly important in the context of dynamic models with export investments. I then review recent evidence that shows PTAs continue to serve an important market access role even if current tariffs and NTBs were zero. The evidence suggests that certain PTAs can credibly secure market access relative to that obtained in the context of WTO and thus serve as insurance against trade wars during large crisis. The trade elasticity with respect to uncertain preferences on the other hand is negligible, which can partly explain the heterogeneous trade effects of PTAs.

Another insight from [Section 4](#) is that certain important dimensions of deeper trade policy cooperation are measurable and contain sufficient variation to identify interesting impacts of PTAs.^a Doing so helps bridge the extreme gap between most of the current theory (and quantitative work), which models only applied tariff changes and constant trade elasticity, and the empirical research that estimates average treatment effects using a PTA dummy but leaves the channel unspecified.

A substantial fraction of trade takes the form of intermediate goods. Moreover, one stated reason for PTAs is to allow members to reorganize the production process across countries more efficiently. In [Section 4](#), I discuss recent empirical work on PTAs where intermediate good linkages can generate additional trade effects relative to the traditional view that focuses on final goods. This occurs for example due to multiple border crossings, which translate into higher trade elasticities when protection is low.

^a For a review of the approaches and estimates of the impacts of trade policy, see [Goldberg and Pavcnik \(2016\)](#).

In [Section 5](#), I address two questions. First, what are the motives for PTAs and the evidence for the mechanisms underlying them? Second, what are the empirical determinants of the formation of PTAs and their policies? I start in [Section 5.1](#) by reviewing the standard trade off in the context of traditional PTAs and the evidence on the mechanisms behind them: trade creation, diversion, and terms-of-trade effects. I then describe some nontraditional motives for PTAs. These motives reflect political economy considerations and international bargaining externalities, as well as some provisions in PTAs documented in [Section 2](#), both economic (eg, FDI, technology diffusion) and noneconomic (eg, environment, human rights, conflict, democracy). I describe the still scant evidence for some mechanisms underlying these nontraditional motives.

In [Section 5.2](#), I review the empirical determinants of (i) PTAs between pairs of countries and (ii) endogenous preferential tariff levels. The potential for bilateral trade plays an important role in the probability of PTA formation, which confirms the importance of addressing endogeneity in gravity estimates. There is suggestive evidence that trade diversion also plays a role but causality is not yet established; this and other aspects of the determinants of PTAs remain fertile ground for research. One promising avenue is to explore preferential tariffs and other product level policy data. This may allow us to test sharper predictions, establish causal effects, and identify certain structural parameters that may be used to quantify interesting counterfactuals.

In [Fig. 1](#), we see not only that PTAs continued to proliferate after the creation of the WTO but also that no major multilateral trade negotiation has succeeded since. The Doha Round was launched 6 years after the creation of the WTO and it is yet to be concluded. WTO membership has continued to expand and this along with the expansion of PTAs implies that a large fraction of trade between WTO members is between preferential groups. The fraction of country pairs in the WTO that also belong to PTAs increased by a factor of 10 in that period and in 2010 they accounted for over 50% of trade between WTO members, even if not all is done under preferential tariffs. This raises the question of how preferential and multilateral agreements and policies interact, which I analyze in [Section 6](#).

A similar interdependence question arose in the early 1990s when PTAs started to proliferate while the Uruguay Round (UR) stalled. This generated a number of important theoretical insights. Some have implications for the equilibrium structure of agreements, which are hard to test empirically. An alternative approach focuses on estimating the implications of the theory for preferential tariffs and how they change the incentives to apply tariffs against nonmembers. I conclude that the existing empirical research has provided important insights on the effects of preferences on protection against nonmembers. This should be complemented with further analysis of the incentives to change deeper policy cooperation, eg, uncertainty and NTBs. These deeper policy dimensions are increasingly important determinants of trade and thus of the potential of PTAs to affect nonmembers.

The long-standing importance of PTAs in the trading system has generated a number of important contributions that review them. [Baldwin and Venables \(1995\)](#) provide a comprehensive analysis of the allocation, accumulation, and location effects of regional

integration. Some of their insights from economic geography models are still relevant and I will not attempt to update them. They also discuss some systematic implications of PTAs for the multilateral trading system but since then there have been considerable theoretical and empirical advances—some reviewed by Freund and Ornelas (2010) and also in this chapter. Krishna (2008) reviews the theoretical literature focusing on static impacts of PTAs, mostly in a Vinerian setting, which I do not address except to place more recent work in context. WTO (2011) provides interesting analysis on the nature and motives of recent PTAs. Bagwell et al. (forthcoming), Maggi (2014), and Grossman (2016) review the literature on trade agreements more generally with some reference to PTAs as well but do not address some of the core issues in this chapter, such as the trade effects and empirical determinants of PTAs.

There are some important lessons and guidance for future research, which I highlight throughout the chapter and in the final section. Befitting this interesting and important topic the main conclusion is that we have learned much about PTAs from recent research but many interesting questions remain to be addressed using existing and new theoretical, empirical, and quantitative approaches. The Online Appendix (<http://dx.doi.org/10.1016/bs.hescop.2016.04.013>) describes the data and programs available to replicate and extend the empirical analysis.^b

2. STYLIZED FACTS AND A TAXONOMY

We start by defining a PTA and providing some stylized facts about them based on a common classification that emphasizes differences in the extent of trade preferences. We then analyze a rich dataset of characteristics of modern PTAs. We use it to propose two key dimensions along which to classify these agreements: policy depth and breadth, where the latter includes economic and noneconomic issues. We argue that these dimensions are salient and useful in framing and organizing the current analysis discussed in this chapter and in pointing the way for future research.

2.1 A Definition

The nature of PTAs' membership, issue areas, policy coverage, and depth is diverse and rapidly evolving. Therefore, any attempt at a taxonomy requires a broadly defined domain under which existing agreements fit and flexible subdomains to accommodate any new PTAs. Accordingly, I will use the following definition:

A PTA is an international treaty with restrictive membership and including any articles that (i) apply only to its members and (ii) aim to secure or increase their respective market access.

The requirement that some articles of the treaty apply only to the members emphasizes the discriminatory nature of PTAs. The requirement that the *aim* of some of those

^b Available at http://terpconnect.umd.edu/~limao/handbook_pta.

policies is to improve market access for goods or services excludes agreements with policies that affect trade but do not aim to improve market access (eg, the Montreal Protocol and its ban on products containing CFCs; or bilateral agreements focusing solely on investment or intellectual property rights). The definition allows for agreements that include nontrade-related policies provided they also include policies aimed at improving market access for at least one member.

The restrictive membership requirement reflects accession constraints and excludes agreements with open membership rules, such as the WTO.^c The nature of accession restrictions that characterizes PTAs is varied. In some cases, the restrictions are regional (eg, European Union, North American FTA) but increasing numbers of PTAs are transcontinental. In other cases they are related to income, eg, concessions under the generalized system of preferences (GSP) apply only to developing countries. In contrast, any state or independent customs territory is eligible to accede to the WTO (Article XII, WTO) and a similar rule applied under GATT. While eligibility does not guarantee membership in the WTO, no other international treaty that satisfies the rest of the definition above has this type of open membership; and if any were to adopt it then eventually it may no longer be able to discriminate (eg, if it achieves universal membership) and thus it should not be defined as preferential.^d

Our definition encompasses the narrower view of a PTA as an exception to the WTO's Article I—MFN—whereby a WTO member is expected not to discriminate against others (cf. [Bagwell and Staiger, 2002](#)). PTAs are then defined as exceptions to MFN allowed by either Article XXIV, which allows subsets of countries to reciprocally set tariffs below MFN values if they are set to zero on substantially all trade, or the enabling clause, which does not place those restrictions on developing countries.

Our definition encompasses a broader set of agreements that include (i) policies regardless of their MFN status, which is useful because PTAs address policies not negotiated in the WTO and (ii) countries regardless of their GATT/WTO membership, which is useful because that membership is only recently becoming universal and even in 2010 about 10% of trade between PTA pairs included at least one non-WTO member.^e

2.2 Common Classification of PTAs and Their Evolution Over Time

Our definition encompasses the common subclassifications of PTAs proposed by [Frankel et al. \(1997\)](#), which are commonly used to code such agreements, eg, by [Baier et al. \(2014\)](#), into the following mutually exclusive groups:^f

^c There are other differences between the WTO and PTAs but also similarities, which we discuss in the Online Appendix.

^d Some agreements such as APEC have considered but not adopted open membership (cf. [Bergsten, 1996](#)).

^e In the [Online Appendix II](#), we show the evolution of trade shares by PTA and WTO membership. In 2010, trade between pairs of countries with a PTA was 55% if both were WTO members and 5% otherwise.

^f The latest version of that data, which is used below is at www3.nd.edu/~jbergstr/DataEIAs/September2015.

- Nonreciprocal PTAs (NRPTA), providing one way preferential tariffs, eg, the GSP.
- Reciprocal PTAs (RPTA) providing two-way preferences on only part of the trade, eg, the Latin American free trade area started in 1960.
- Free Trade Areas (FTAs) providing two-way preferences and eliminating tariffs on a substantial part of the trade, eg, NAFTA.
- Customs Unions (CU), which are FTAs with common external tariffs, eg, Mercosur.
- Common markets (CM) such as the European Union, which adds free movement of capital and labor to a CU.
- Economic Unions (EU), which are CM with additional monetary and fiscal policy coordination such as the Economic and Monetary Union of Central Africa (1999) and the Euro area countries.

This classification was constructed in a way that suggests increasing economic integration as we progress from nonreciprocal or shallow reciprocal agreements to FTAs and CM. We can then calculate the share of world trade by the bilateral pairs in each type of agreement. Given the small number of agreements in the last three categories I aggregate them into a single group: CU-CM-EU. The share of world trade between pairs of countries that had any of these PTAs rose from around 22% in 1965 to 60% in 2010. The breakdown for each of the four subcategories is shown in Fig. 2 at 5-year intervals. I note the following key points:

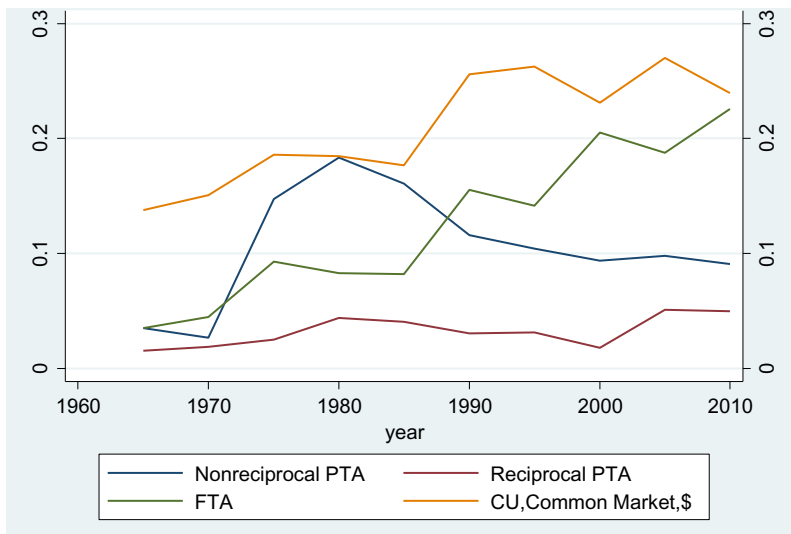


Fig. 2 World trade shares by type of Preferential Agreement.

- NRPTAs' trade increased in the 1970s, partially as a result of Special and Differential treatment (cf. [Ornelas, 2016](#)), declined in the 1980s and stabilized since the 1990s around 10%.^g
- RPTAs' trade is small throughout and its relative importance declines between 1980 and 2000, and increases since then, partly due to an increase in their number.
- FTAs' trade shows a large increase from 3.5% in 1965 to 22.5% in 2010.
- CU–CM–EU has the largest share through the whole period, partly due to the large amount of intra–European Union trade. That share increased from about 14% to 24%. Based on trade shares, it is clear that since the mid–1980s there has been a sharp increase in the relative importance of “deeper” PTAs. This increase is at least partially due to the fact that FTAs have become more prevalent and membership in common markets has expanded. At the start of our sample the European Union included only the 6 founding members, by 1995 it included 15, and then 27 in 2010. More generally, the fraction of country pairs with any PTA that have either an FTA or CU–CM–EU has increased substantially: from about 1/8 in 1985 to 1/3 by 2010.^h

Using this data we can start to classify PTAs along their policy depth and breadth. The trade policy that defines depth in this data is almost solely applied tariffs and whether they are applied only to one member (NRPTA) or both (RPTA); whether that tariff is zero on most goods (FTA) and whether members set a common external tariff (CU). This reflects the traditional Vinerian view of PTAs. These are useful measures of depth but they are increasingly insufficient as MFN tariffs are lowered and the focus switches to other policies. There are many aspects of a CM, some of which I will classify under deeper cooperation (eg, behind-the-border policies that affect trade), and others as broader cooperation, eg, the movement of factors. An increasingly large number of agreements fall between the common classification of an FTA and a CM and sometimes beyond them and thus below I propose augmenting this common classification along the depth and breadth dimensions.

2.3 A Taxonomy of “Modern” PTAs

I now use data on different provisions contained in modern PTAs to provide a more detailed taxonomy and identify the prevalence of certain features that may warrant future research.

The data were originally compiled by [Horn et al. \(2010a\)](#) to analyze the precise content of 28 agreements that either the United States or European Union signed with WTO members up to 2008 under Art. XXIV of the GATT or Art. V of GATS. The data were

^g These values include all exports that the beneficiary country makes to the preference granting country, regardless of whether they obtain a preference so it overestimates the share of preferential trade.

^h We could further summarize the evolution of PTAs by economic development of partners, location, participation in other agreements, colonial or language ties, etc. For these and other interesting facts pertaining to the characterization of PTAs, we refer the reader to [WTO \(2011\)](#). We will examine the determinants of PTA formation in [Section 5](#).

subsequently extended by the WTO to cover 100 agreements including non-WTO members through 2011.ⁱ The 52 types of provisions identified by the authors are listed in Table A1 (<http://dx.doi.org/10.1016/bs.hescop.2016.04.013>) of the Online Appendix. Horn et al. classified the policy areas as already included in the WTO in some form (WTO+) or qualitatively new (WTOX) and then according to their legal enforceability. These are useful classifications given the extensive research on the WTO and can help to highlight what its members may be looking for in PTAs. However, for the purposes of the chapter and future research I have grouped the policy areas into different depth and breadth classifications that I will describe and use.

In generating the taxonomy from this rich set of variables, I was broadly guided by the following considerations. First, the evolution of the GATT/WTO in terms of policy depth and breadth described in detail in the Online Appendix II; second, my own reading from this data and other information about important developments in recent PTAs; third, the economic similarity and relevance of different dimensions; and fourth, an attempt to encompass the traditional classification of PTAs and to extend it in terms of finer depth categories and broader policy cooperation.

An agreement is defined as a PTA if a subset of its policies affects the market access of either country. So the first step, and main focus below, is to argue for the economic relevance of each level of policy depth and breadth in affecting market access.

2.3.1 Depth in Economic Policy Cooperation

The most obvious element of policy depth that affects market access is tariffs—the lower they are the deeper the level of bilateral economic cooperation. This is the reasoning behind separating RPTA and FTAs. While trade agreements typically start by addressing applied tariffs, they eventually tackle other policies (as was the case with GATT). To characterize the depth of cooperation we divide policies into four groups, $\pi^d = \{\tau, \nu, \gamma, o\}$, where

- τ import tariffs: applied or bindings
- ν NTBs
 - contingent protection: eg, antidumping, countervailing measures, export taxes;
 - other: product standards, customs procedures, and other technical barriers.
- γ policies behind-the-border not included in ν that may invalidate national treatment.
 - state aid, procurement, competition policy.
- o other policies that *may* affect market access but can also have direct effects
 - regional, industrial, agricultural cooperation, financial assistance.

Is there a meaningful way in which these policies can be ranked in terms of depth of cooperation? In some cases the answer is clearly yes. Cooperation is deeper in an

ⁱ The data with the list of agreements are available at https://www.wto.org/english/res_e/booksp_e/anrep_e/wtr11-anatomy_ptas_e.xls. Dür et al. (2014) provide an even more comprehensive coverage of agreements, including their features and sector coverage.

agreement that not only specifies a zero tariff today but also binds it at a maximum level forever. In other cases it is less obvious so we now provide some discussion of this point.

How prevalent are the NTBs included in ν and do they represent deeper cooperation? As tariffs fall, other barriers become more salient, so if PTAs also tackle certain NTBs, then it is reasonable to claim they involve deeper policy cooperation. Among the easiest of the NTBs to identify and address are customs procedures for shipment inspection, which can explain why 93% of the PTAs in the sample explicitly addressed it. Others, such as different types of contingent protection, may not be as simple. Still 86% of PTAs address antidumping in some way and 72% address the use of duties to countervail subsidies.

Countries can also use product standards as nontariff barriers to trade. Countries often have different product standards, eg, criteria for a particular classification of a good, health safety requirements, etc. Some are designed for consumer protection but they can also be used to discriminate against foreign producers either explicitly or implicitly (eg, by requiring additional testing). Countries can thus decide to harmonize or mutually recognize their standards.^j Doing so is harder than reducing tariffs both because different standards can reflect different preferences (eg, genetically modified food in the United States vs European Union) and because it is harder to verify if a standard is enforced differently on domestic and foreign goods. This may explain why only 60% of PTAs in 2011 addressed any standards.

After a foreign product crosses the border, government policy may still treat it differently from national products. One obvious policy would be differential enforcement of a product standard, which we include in ν . In γ , we include other policies that may also impair national treatment and market access. Some of these reduce expected trade costs, eg, establishing/maintaining nondiscriminatory competition policy, possibly by an independent authority, as done in 63% of PTAs. Others increase expected demand by providing information on state aid to enterprises (59%) and regulating public procurement (53%). Some of these provisions may also reflect a commitment motive by governments attempting to address time-consistency or political economy constraints, as we discuss in [Section 5.1](#).

Cooperation along other dimensions may also affect market access or government objectives directly. We include these in the “other” category (eg, regional, industrial and agricultural cooperation, and financial assistance). Their prevalence is still relatively low ranging from 1% to 32% of PTAs depending on the policy.

To summarize the current prevalence of each of these four groups of policies we need a metric to aggregate the categories. In the first column of [Table 1](#), rows 1–4, we ask what fraction of PTAs address tariffs and at least one of the policies in each of the other

^j This can reflect both gains from cooperation (reduce protection substitution) and coordination (minimize duplication of costs to meet similar standards in the markets). See [Ederington and Ruta \(2016\)](#) for a detailed analysis.

Table 1 Depth and breadth in economic policy cooperation

	Breadth					
	Type of trade		Technology	Factors of production		
	Goods	Goods and services	Innovat./IP	Capital	Labor	All
Depth						
Tariffs (τ)	1.00	0.63	0.79	0.76	0.58	0.42
τ and nontariff barriers (ν)	0.98	0.64	0.81	0.78	0.59	0.43
τ and behind border policies (γ)	0.89	0.66	0.83	0.78	0.62	0.46
τ and other policies (ρ)	0.60	0.70	0.92	0.92	0.77	0.53
All (τ , ν , γ , and ρ)	0.56	0.73	0.93	0.91	0.79	0.55

Notes: Share of agreements that address different combinations of depth and breadth.

Author's calculations from data in [WTO \(2011\)](#).

categories. So, 100% of the PTAs address at least one tariff-related policy, 98% address tariffs and *any* of the policies in ν , 89% address tariffs and any of the policies in γ , and 60% tariffs and policies in ρ . The last row shows that 56% of PTAs addressed tariffs and at least one of the policies in each and every one of the subgroups.^k

Within each of the policy groups there is an additional dimension of depth worth noting: whether the PTA addresses only the current or also the expected policy. Certain PTAs not only reduce current tariffs but they also set them to zero in all goods, which may reduce uncertainty about future protection. Constraints on certain NTBs, eg, contingent protection, may also be motivated by concerns with reducing future policy uncertainty. The same can be said of procurement and competition policy since the agreements do not stipulate that specific current government procurement contracts must be directed to a member but rather set out rules for how to address such situations in the future. The impact of deeper PTAs that can credibly reduce uncertainty about future policy is examined in [Section 4](#).

2.3.2 Breadth in Economic Policy Cooperation

To span the breadth of economic policy cooperation I divide policies by their impacts on:

- Type of trade: Goods and services.
- Technology: Innovation and diffusion, Intellectual property.
- Factors of production: Investment/capital and labor.

^k The reader can use the online data to calculate the average prevalence of other groups of policies in PTAs.

PTAs have historically liberalized trade in goods first, as did the GATT/WTO.¹ But currently both the WTO and 63% of PTAs include provisions for expanding market access in services.

Broadening PTAs to include trade in services has clear implications for market access. The implications of including policies relating to technology and the factors of production are less obvious. But it is conceivable that they affect market access, and since they can also have other direct effects on members' economies we believe they are important dimensions of the breadth of cooperation.

About 79% of PTAs include some provision regarding technology, which I divide into intellectual property and innovation/diffusion. Among the prominent examples of the latter are provisions that promote technology transfer; joint research projects; exchange of researchers and development of public–private partnerships. About 43% of the PTAs include at least some innovation and diffusion provision so one could explore this variation to examine if those agreements tend to generate stronger dynamic gains in the form of increases in R&D and/or productivity. Intellectual property right protection clauses are now part of 61% of PTAs either in a form similar to the WTO or beyond it.^m

The factor of production categories capture either investment or labor provisions. Currently, 76% of PTAs include at least one provision related to investment such as (i) the liberalization of capital movement and prohibition of new restrictions (58%) and (ii) requirements for local content and export performance of FDI (45%).

Only 58% of PTAs address labor market provisions, which include any of the following: illegal immigration (9%) or labor market regulations, visa and asylum, coordination of social security, each included in around 25% of PTAs.

The final column of the table shows that 42% of PTAs in 2011 addressed at least one of the policies in each one of the groups: goods, services, technology, investment, and labor.

2.3.3 Complementarity of Depth and Economic Breadth

An interesting question related to our taxonomy is how economic depth and breadth interact. In particular, is policy cooperation depth as we define it shallower if the agreement is broader? This is possible if negotiation and enforcement resources are scarce for example. The data suggests that the opposite is true. The probability that an agreement addresses services issues increases steadily from 63% as we move down to include additional depth in policies, all the way to 73%. This complementarity is present for each of the breadth categories. Overall the unconditional probability of a broad agreement—one with clauses on trade in goods and services, technology and each factor of production—is 42% but it increases to 55% if we condition on it being deep.

¹ In specific agreements, eg, the EEC, we also know that industrial goods are covered before agriculture.

^m The impact of IPR on market access as measured in terms of export value is typically ambiguous. But from the perspective of a developed exporting country with large number of firms with patented products, IPR protection in the foreign market is valuable. For a detailed discussion of IPR, see [Saggi \(2016\)](#).

It would be interesting to have a theory of the design and evolution of cooperation in agreements that explain the interaction between these dimensions of economic depth and breadth.

2.3.4 Breadth in Noneconomic Policy Cooperation

Using this data, we also obtain the prevalence in PTAs of the following “noneconomic” issues:

- Environmental laws (46%) and health (10%).
- Human rights (14%) and political dialog (14%).
- Illicit drugs (13%), money laundering (12%), and terrorism (6%).

The exact description of each of these categories is provided in Table A1 (<http://dx.doi.org/10.1016/bs.hescop.2016.04.013>) where we also show whether those provisions are legally enforceable.¹¹ Provisions on environmental laws are enforceable in 35% of the cases where they are included but those on human rights, health, or illicit drugs are not. Moreover, the United States is involved in all the agreements that have legally enforceable environmental provisions. In contrast to these, provisions on the movement of capital and investment are legally enforceable in 98% of agreements where they are included. Whether a clause is legally enforceable is not a necessary or sufficient condition in determining if the agreement affects the relevant issue. So, given the potential importance of these noneconomic issues, future research should examine if they are affected by PTAs.

In contrast to the complementarity we observed between depth and economic breadth, we find some substitutability between depth and noneconomic breadth. The unconditional probability of a PTA including noneconomic areas in 2011 is 56% but it drops to 23% if we condition on policy being deepest (ie, on an agreement addressing tariffs, nontariff, government policies on procurement and other). In [Section 5.1](#), we discuss how issues with nonpecuniary externalities, such as the ones above, may provide a rationale for trade preferences.

2.4 Trends in Modern PTAs

We provided a snapshot of PTAs in 2011. In the Online Appendix II, we graph the evolution of the share of PTAs with different provisions since 1991. Here we point to a few key points. In terms of depth the most important trend is the increased prevalence of provisions addressing contingent protection, product standards, and public procurement. In terms of economic breadth, since 2000 there has been an increase in service provisions and labor market regulations. Finally, in terms of noneconomic areas we see increases in environmental laws from less than 30% to almost 50% and also in human rights and illicit drugs.

¹¹ For example, environmental laws include: development of standards; enforcement of national laws; establishment of sanctions for violation of environmental laws; publications of laws and regulation.

2.4.1 Emerging Dimensions and Complementary Data

The data we explore in this section are rich and allow for an improved taxonomy relative to the common classification. Our understanding of PTAs will be further improved if the data are extended to more PTAs and complemented with information about current and emerging important dimensions in PTAs.

First, we can complement this with data on common currency to include an additional dimension of breadth. We can also complement the depth dimension and further refine the tariff classification according to whether the agreement is reciprocal, an FTA or CU.

Second, the policy depth of trade policy focuses on measures affecting final goods. Going forward, it is important to collect data on policy related to trade in intermediate goods. The continuing slicing of the production chain to take advantage of economies of scale and/or cost differentials is a potentially important motive for PTAs. Thus incorporating policies that affect trade in intermediates is important. Some such data are available: rules of origin can affect whether a good is eligible for a preference and variation in such rules (such as ability to accumulate value added shares across members) can be used to explore the impact of these agreements. We return to this in [Section 4](#).^o

Additional information would also be useful on policies that affect multinational investment and incentives for arms-length trade. These include any rules on transfer pricing and taxation of profits as well as on investment dispute systems, which are controversial items proposed in the TTP and TTIP.

2.5 Stylized Facts

In this section, we provided a few stylized facts, some of which help guide subsequent analysis, namely:

- (1) The large and growing trade share of bilateral “deeper” PTAs in world trade. The growing number of bilateral PTAs has translated into an increase in the share of world trade between their members from 22% in 1965 to 60% in 2010. Moreover, FTAs and customs and economic unions have become relatively more important, particularly since the mid-1980s.
- (2) The large and growing interdependence across agreements. The share of WTO country pairs with PTAs rose from 2% in 1965 to over 25% in 2010 and their corresponding trade share within the WTO rose from 30% to 60% in that period. Moreover, there has been an increase in the overlap of PTA memberships for any given country.

^o Other instruments may not yet be used but could in the future. For example, the increased data on value added trade may eventually lead governments to consider charging tariffs on net rather than gross value.

- (3) The deep, broad, and heterogeneous cooperation in modern PTAs. Most recent PTAs go far beyond applied tariff reductions on goods. Policy depth tends to be higher in agreements with broader cooperation in economic issues but not in non-economic ones. Moreover, there is heterogeneity across and within PTAs in the legal enforceability of different provisions.

Fact 1, along with the primacy of preferential market access in the definition of PTAs, explains the main focus in [Section 3](#): to identify if PTAs increase bilateral market access. Most research focuses on identifying an average trade effect, but Fact 1 suggests this effect is heterogeneous across agreements. The traditional/Vinerian view of PTAs predicts that any trade effects and any heterogeneity in these is caused by differences in applied tariffs, so [Section 3](#) also examines if these can explain the full trade effect of PTAs or its heterogeneity.

Fact 3 points to the many other policy dimensions in modern PTAs. In [Section 4](#), we study how some of the deeper policy cooperation affects firm decisions to trade. In [Section 5](#), we discuss some evidence of the effects of PTAs on certain nontrade outcomes. The taxonomy I provide may also help group provisions to characterize why PTA have heterogeneous trade effects, as we discuss in [Section 3](#). The rich set of provisions and their enforceability may also be useful in future research to test sharper predictions about the effects of PTAs on outcomes beyond aggregate trade including innovation, FDI and labor outcomes.

Fact 2 provides one motivation for studying the determinants of the formation of PTAs in [Section 5](#) and the interdependence between agreements in [Section 6](#).

This section highlights one additional feature worthy of future research: the expansion and evolution of cooperation in trade agreements. Both multilateral agreements and PTAs have experienced significant increases in membership, policy depth, and breadth. To my knowledge there is no theory of the evolution of cooperation along these dimensions, eg, why are applied tariffs tackled first, then NTBs and eventually broader policies. Such a theory could help us understand (i) the limits of cooperation within the WTO and thus some of the incentives to form PTAs as well as their key features, which could also help explain facts 2 and 3.^P

3. TRADE AND WELFARE EFFECTS ON MEMBERS

A central feature of a PTA is that certain policies aim to increase market access between its members. Therefore, our starting point is to examine estimates of the trade effects of PTAs on their members. I start with ex post estimates of bilateral PTAs that rely on

^P One interesting path is to model agreements as endogenously incomplete contracts as done by [Horn et al. \(2010a,b\)](#).

gravity estimation and highlight two apparent puzzles generated by them. First, why did PTA formation explode in the last 20 years when until the early 2000s the estimates of trade effects were very heterogeneous and in some cases modest—the PTA formation puzzle.⁹ I discuss how the econometric resolution of this puzzle generated long-run estimates that are in some cases so large that they raise a different one: whether PTAs can create so much bilateral trade given the observed levels of trade barriers—what I refer to as the PTA trade elasticity puzzle.

I discuss evidence that the PTA effect is particularly large for a certain class of deeper and broader agreements, suggesting it captures more than a standard static tariff reduction. Novel evidence indicates that controlling for bilateral tariff changes only reduces PTA effects slightly and that PTAs themselves have only a modest effect in reducing applied tariffs for member countries. I then discuss additional mechanisms that may be able to explain the PTA trade elasticity puzzle. [Section 4](#) examines some of these mechanisms in detail.

I provide gravity estimates for PTAs in 1965–2010 within a unified dataset where we can examine the importance of alternative econometric approaches and assumptions. I do so while controlling for WTO membership and also discuss the WTO estimates in the literature, which are interesting in their own right and also a reference point for the effects of PTAs. The data and programs are available for the reader to explore and test alternative hypothesis.

I conclude the section by briefly reviewing structural general equilibrium estimates of PTAs on trade and welfare of its members from new quantitative trade models and traditional CGE applications.

3.1 Ex Post Trade Effects: Naïve Gravity and a PTA Formation “Puzzle”

3.1.1 Gravity Approach

Since [Tinbergen \(1962\)](#) the gravity equation has been used to estimate the effects of bilateral PTAs on aggregate trade between its members. The empirical success of the gravity approach and its theoretical underpinnings make it a potentially useful tool for this purpose ... provided it is correctly used. To interpret the evolution of PTA estimates we start with the formulation of a gravity equation in a cross section[†];

$$T_{xm} = aX_x M_m \phi_{xm} \quad (1)$$

where T_{xm} represents the value of exports from x to importer m ; X_x captures the supply “capabilities” of x to any m ; M_m captures any importer market characteristics that affect its

⁹ This is only a puzzle if the objective of policy makers is market access and they are influenced by academic research!

[†] This corresponds to definition 1 in [Head and Mayer \(2014\)](#), which I follow for ease of comparison.

demand from any x . The term $\phi_{xm} \in [0, 1]$ is a general measure of bilateral market access; a is a constant.

Before discussing how gravity has been used and the interpretation of different estimates, we define what effects we are interested in. Let π_{xm}^{PTA} denote the set of potential policies that change depending on whether $PTA_{xm} = \{0, 1\}$. Depending on the set of policies and economic structure we would have a particular mapping from π_{xm}^{PTA} to the equilibrium value of each variable on the right side of (1) (more on this below). We define the general trade effect of a PTA between x and m as

$$\widehat{T}_{xm}^{ge} = \widehat{X}_x + \widehat{M}_m + \widehat{\phi}_{xm} \quad (2)$$

where $\widehat{M}_m \equiv \ln M_m^1 / M_m^0$ and M_m^{PTA} represents the equilibrium value depending on the existence of PTA; similarly for the remaining variables. The PTA can affect policies that enter directly into X and M and thus affect all partners regardless of membership but here we focus on bilateral policies that apply differentially to members. So (2) represents the total growth in member exports caused by their PTA, while holding any other policies constant but, allowing for any general equilibrium effects of bilateral policies that can magnify or counteract the partial trade effect of a PTA via the bilateral market access function, which we define as:

$$\widehat{T}_{xm} = \widehat{\phi}_{xm} \quad (3)$$

Nearly all ex post gravity estimates attempt to estimate either this partial effect for a specific agreement or its average across agreements. If we believe the PTA is essentially a reduction in bilateral costs then the partial effect should be positive regardless of the exact economic structure that generates the gravity structure. That structure will be important in determining the general effect, which we discuss at the end of the section.

3.1.2 Average Partial Effects and Selection Bias

The structure of gravity equations is conducive to estimating the average partial effect of PTAs as an average treatment effect (cf. [Baier and Bergstrand, 2007](#)). Doing so helps unify our understanding of problems with earlier estimates as different sources of selection bias.

What are the basic conditions for obtaining a consistent estimate of the average partial effect over all PTAs, $\widehat{\phi}$? To answer this and identify the different sources of bias we first note the following property of gravity models. The bilateral market access between any xm pair has no *direct* effect on exports from the rest of the world to m (or any other countries) if we condition on X_x and M_m . This suggests we can estimate $\widehat{\phi}$ as an average treatment effect under the standard condition that the treatment, PTA membership, satisfies the conditional independence assumption. That is, conditional on (i) the relevant additive determinants of the exporter and importer characteristics and (ii) the determinants of bilateral access, the PTA treatment is “random” so we can obtain \widehat{T} as the conditional average difference in $\ln T$ between pairs of countries in a PTA and those outside.

While the conditional independence cannot be directly tested, we can examine how the results change as we move toward meeting it in terms of the econometric and economic model. To do so let us write the true value of the (log) X_x , M_m , and ϕ_{xm} each as a (different) function of some vector of observable variables, summarized by $Z_i = \{Z_x, Z_m, Z_{xm}\}$, and an error term for each $\{\epsilon_x, \epsilon_m, \epsilon_{xm}\}$. Using this we write log exports in its conditional expectation form, where u_{xm} is a random error.

$$\ln T_{xm} = \hat{\phi} PTA_{xm} + Z_i' \beta + \epsilon_x + \epsilon_m + \epsilon_{xm} + u_{xm} \tag{4}$$

Most gravity estimates adopt some form of (4) with different control variables or fixed effects, for different country and time samples. Moreover, the large majority adopts ordinary least squares in which case the estimated parameter on the PTA variable, $\tilde{\phi}$, is equal to the average partial effect and a PTA selection bias term: the difference across members and nonmembers of the error terms for exporter capabilities, importer demand characteristics, and bilateral market access.

$$\tilde{\phi} \equiv \underbrace{E[\ln T_{xm} | Z_i, 1] - E[\ln T_{xm} | Z_i, 0]}_{\text{Estimated PTA effect}} = \underbrace{\hat{\phi}}_{\text{Partial effect}} + \underbrace{E[\epsilon_i | Z_i, 1] - E[\epsilon_i | Z_i, 0]}_{\text{PTA selection bias}} \tag{5}$$

where $\epsilon_i \equiv \epsilon_x + \epsilon_m + \epsilon_{xm}$. I wrote the partial effect as an average over all PTAs but several studies focus on specific PTAs. This is reasonable since some PTAs may be deeper or broader than others. The approach in (5) extends to these cases, if we take the expectation over each of the relevant PTAs, but at the potential cost of small sample bias.

In implementing this approach we require a random sample from the population over which we want to estimate the effect. I would argue that the relevant population is typically bilateral pairs with positive trade in the period under study since countries with no trade generally have no PTA and are unlikely to gain from and thus ever forming one. If positive trade pairs are the relevant population then the only sample selection issue is due to the availability of covariates. If the population were any bilateral pair then there would be a selection bias due to the omission of zeros in the log linear approach, which we will discuss as well. In sum, the discussion below will refer to selection bias arising from three potential sources

- (1) Small PTA sample size.
- (2) Choice of controls and estimation approach.
- (3) Sample selection.

3.1.3 Naïve Gravity

I now discuss estimates that follow what [Head and Mayer \(2014\)](#) define as the naïve gravity approach, which is characterized by using GDP to proxy for X_x and M_m . The early literature is too long to summarize and we now understand that much of it suffers from basic estimation problems. Therefore, I just highlight the heterogeneity and fragility of some results and how they are subject to different types of selection bias. Section 3.2 evaluates the importance and solution of each source of bias.

Until recently most gravity estimates used GDP to proxy for X_x and M_m and included bilateral distance and a PTA dummy as part of the bilateral market access. The remaining controls vary across different studies. The results were mixed. Frankel et al. (1997) consider a number of different agreements and time periods. He finds small or insignificant effects for the EC in several time periods (and reports similar results by others) but strong effects for agreements such as Mercosur (p. 62).

The estimates found were heterogeneous across the type of agreement examined, the time period, country sample, and controls. So much so that different authors could, and did, cite the same paper for finding either a strong effect or none for the same PTA.[§] In the presence of numerous and sometimes contradictory estimates, meta-analysis can provide a useful measure of their robustness and variability. World Bank (2005) reports the results from 362 estimates of $\tilde{\phi}$ in 17 studies that span different agreements, time periods, and specifications. One-third of the estimates is statistically insignificant, 12% are negative and significant, and only 54% are positive and significant. The mean estimate is 0.79 but the standard error is 1.3.[†]

3.1.3.1 Small PTA Sample Size

Many of the studies suffer from a small PTA sample, as they rely on cross-sectional data or use data available only through the mid-1990s. In particular, these studies either focus on a single agreement or attempt to estimate separate effects for multiple ones using aggregate trade, often defined as the sum of imports and exports, so there is a single observation per country pair in a given year.

3.1.3.2 Choice of Controls

The second issue with earlier estimates is their reliance on proxies for exporter and importer characteristics that fail to account for systematic determinants of trade.

[§] In explaining why no results are found for the EC Frankel et al. (1997, p. 87) notes that Bergstrand (1985) also finds none. Ghosh and Yamarik (2004, p. 370) claim that Bergstrand (1985) provides evidence that “European trade blocs increased trade during the 1960s and 1970s” and cite it as an example of an emerging consensus that PTAs are trade creating, which they then go on to challenge. Baier and Bergstrand (2007) claim that Bergstrand (1985) finds an insignificant effect for the EC and proceed to show how their approach provides larger robust estimates. The results in table 2 of Bergstrand (1985) show significant effects for EFTA in the 60s and 70s and for EC in 60s but not in 70s. It also shows the number of observations in each cross section: 210, because the study uses total trade by only 15 developed countries.

[†] The study also reports the average over 254 estimates that attempt to estimate an “overall” impact of PTAs: the sum of the member effect and any trade diversion effect with nonmembers. In this case, 42% of the overall impact estimates are negative and significant and only 34% are positive and significant. To do so those studies include a PTA variable equal to one if x is a member of any PTA, which after controlling for PTA_{xm} , is meant to capture if there is less trade with nonmembers. The sum of these effects is the overall impact, which as we will see is not identified once we fully control for the exporter and importer characteristics. We discuss trade diversion in Section 5.

To see the importance of this point consider the core determinants generally used: distance, importer, and exporter GDP. Suppose that a specific group of countries in a region is pursuing unilateral liberalization, for instance Brazil and other Latin American countries in the early 1990s. This unilateral liberalization implies their average trade is higher than predicted by the standard gravity variables and the same is true of their bilateral trade. If those countries also have a PTA, say Mercosur, then the earlier gravity studies would estimate a large effect. Clearly there is an omitted variable, unilateral liberalization, which can be a source of selection bias, essentially the exporter and importer errors in (5) are higher for certain PTA members. Conversely, those errors may be lower for countries that have already liberalized unilaterally or have trade below what they would like or expect based on their characteristics and are trying to boost it via PTAs.

Partly as a recognition of this potential problem, several researchers used a large set of covariates that include income per capita, geographic features, historic ties (colony, language), exchange rate measures (float vs fixed, volatility, currency unions), and in one case aggregate trade policy. The recognition that gravity could also be derived from comparative advantage models also led some to include factor endowments. However, more is not always better. There are three potential pitfalls of using multiple typically atheoretical controls.

The first potential pitfall is the possibility that one can find at least a combination of covariates, accidentally or by design, that is particularly good at explaining trade for a particular set of countries. Ghosh and Yamarik (2004) argue that this can lead researchers to report only the results that accord with their own priors. To analyze this systematically they perform an extreme bound analysis. They take data for 186 countries and six 5-year periods from 1970 to 1995 and regress trade on the core variables: GDPs and distance as well as PTA effects and time dummies. They then consider the estimated effects of PTAs under different combinations of 16 other possible determinants. When all these determinants are included they find positive and significant effects for 8 out of 12 PTAs and negative effects for the European Union and NAFTA (the latter is significant). When they calculate the extreme bounds of those estimates under alternative combinations of covariates they find that the range for each of the PTAs always includes zero and is extremely wide, eg, $\tilde{\phi}_{EU} \in (-2.2, 2.5)$.^u

The second concern with multiple atheoretical controls, is the interpretation of the resulting estimate. Suppose that PTAs affect trade solely via changes in trade policy and that we control for a country-specific trade policy openness index. We may then find that

^u Eicher et al. (2012) argue this type of analysis is too extreme and argue instead for Bayesian model averaging, which weights the models according to their ex post explanatory power, whereas extreme bound analysis provides equal weight to all models. Their Bayesian approach generates results closer to the OLS estimates.

the PTA has no impact, which can either be due to the fact that the index captured the PTA policy change or the fact that the country underwent unilateral liberalization.^v

3.1.3.3 Sample Selection

The third type of selection suffered by early econometric exercises relates to sample selection; a problem that can be exacerbated by the use of multiple atheoretical controls. For example, in [Ghosh and Yamarik \(2004\)](#) there are over 31,000 observations with positive trade and information for the core variables but less than half of it is used because of missing data on the multiple covariates. Data collection and quality are not random across countries and can thus result in selection bias. Moreover, this and most other initial estimates use only positive trade, which in this case means discarding 2/3 of the potential country pairs. This may not be as severe an issue if we believe that most countries that never trade would also never have a PTA and thus our target population excludes them.

In sum, not controlling for key determinants of trade generates a selection bias, which can be particularly problematic in a small PTA sample setting. However, multiple combinations of atheoretical determinants are not necessarily an improvement. In fact, the conditional independence assumption of PTAs may hold with one set of covariates but not another, particularly if they include outcomes that are affected by the PTA^w and affect the sample.

3.2 Ex Post Trade Effects: Theory Consistent Estimates

The preceding discussion suggests that we require:

- (1) a clearer theoretical guidance of trade determinants in gravity models and a more robust way to control for them;
- (2) explicit mechanisms for formation and impacts of a PTA to better justify conditional independence; and
- (3) larger PTA samples, either by estimating average effects over similar agreements or using more disaggregated data.

Several of the recent developments in gravity estimation and their application to PTAs address these issues, which we now discuss in turn.

3.2.1 Structural Gravity and Multilateral Trade Determinants

The general formulation of gravity in (1) arises from a variety of different models (cf. [Costinot and Rodriguez-Clare, 2014](#)). Those models can provide alternative interpretations for the multilateral trade determinants, X_x and M_m , but regardless of their

^v A more subtle example is if PTAs affect trade by affecting bilateral exchange rates (eg, the volatility or whether they subsequently form a currency union) then controlling for the latter will affect our estimate of the former.

^w [Imbens \(2004\)](#) makes this point more generally for estimation of average treatment effects.

microfoundation, we can account for them in an econometrically robust way by using importer- and exporter-specific effects. If in addition we specify a bilateral market access function with constant trade elasticities, β_T (across countries) then we do not require any additional structure to obtain estimates of the partial PTA effect where the multilateral terms are consistent with theory. Doing so implies estimating

$$\ln T_{xm} = \hat{\phi} PTA_{xm} + Z'_{xm} \beta_T + \alpha_x + \alpha_m + \epsilon_{xm} + u_{xm} \tag{6}$$

where Z'_{xm} can include distance and other structural determinants of bilateral trade frictions. All importer- and exporter-specific variables are captured by their respective fixed effects, α_m and α_x , which eliminates any country-specific sources of selection bias in (5).

We now provide a structural interpretation of α_m and α_x and explain a potential source of bias when instead of controlling for these effects we use a proxy such as GDP. To do so consider the following Monte Carlo experiment in [Head and Mayer \(2014\)](#). They generate data from a gravity consistent framework and impose $\ln \phi = 0.5 PTA_{xm} - Dist_{xm} + \epsilon_{xm}$, so the true partial effect assumed is $\hat{\phi} = 0.5$; ϵ_{xm} is the only source of error in the data generating process and has a normal distribution. Using the approach in (6) with distance and fixed effects as controls they recover the true partial effect of 0.50 whereas the OLS estimate obtained from replacing the fixed effects with GDPs is only 0.28.

To interpret the downward bias in the estimates without fixed effects we need to understand the theoretical model used to generate the data. In it the exporter and importer characteristics are given by $X_x = Y_x / \Omega_x$ and $M_m = T_m / \Phi_m$ where $Y_x \equiv \sum_m T_{xm}$ represents total value of production and thus income for the exporter; and $T_m \equiv \sum_x T_{xm}$ the total expenditure by the importer. These terms are adjusted by the multilateral resistance terms (cf. [Anderson and Van Wincoop, 2003](#)), $\Omega_x \equiv \sum_j \phi_{jx} \frac{T_j}{\Phi_j}$ and $\Phi_m \equiv \sum_j \phi_{jm} \frac{X_j}{\Omega_j}$, which are expenditure weighted measures of average market access.^x Therefore when we use OLS and GDP the error will reflect the importer and exporter multilateral resistance terms, $-\ln \left(\sum_j \phi_{jx} \frac{T_j}{\Phi_j} \right) \left(\sum_j \phi_{jm} \frac{X_j}{\Omega_j} \right)$. All else equal this term is lower for the PTA countries (since ϕ_{xm} and ϕ_{mx} are higher), which offsets the true partial effect. One interpretation is that the PTA lowers the importers' price index, which reduces its imports from other sources and the OLS estimates will capture the net effect.

In practice, the country effects can also capture other factors that generate selection biases in different directions. For example, if a unilateral liberalization coincided with the PTA then the fixed effect estimates should be lower. On the other hand, if the PTA led to

^x As [Head and Mayer \(2014\)](#) show this derivation requires only that expenditure shares over different export sources are independent of income and uses market clearing conditions.

higher protection against nonmembers than the fixed effects estimates would be higher. The net effect likely depends on the sample. What do the data say?

In [Table 2](#), we employ the data used in [Section 2](#) and estimate partial PTA effects using different approaches on data from 1965 to 2010 at 5-year intervals. Column 1 uses the OLS naïve approach controlling for GDPs and their deflators and year effects whereas column 2 controls for country-by-period fixed effects; both columns 1 and 2 control for common border and language. In this sample the estimated PTA partial effect is smaller after we control for country-by-period effects.^y

It is important to understand the interpretation of the PTA coefficient after accounting for multilateral resistance in this way. The average change in the exports of x to all its trading partners after a PTA is reflected in its multilateral resistance term. Thus the PTA effect reflects only the differential trade with the PTA member relative to nonmembers, so it can reflect creation of trade with this member or diversion away from nonmembers. We may compute the effect of PTAs on nonmembers by combining the estimates with a specific theoretical model, as we discuss at the end of this section.

3.2.2 Endogenous PTAs and Bilateral Trade Determinants

The structural gravity approach in [\(6\)](#) is still potentially subject to selection bias arising from any systematic differences in unmodeled bilateral trade determinants between PTA and non-PTA pairs. Let us consider what these may be and how they have been addressed.

Consider applying OLS to [\(6\)](#) in a cross section with a PTA indicator but no other bilateral determinants. It should be clear that the partial effect estimated is identified only from countries that have PTAs and is equal to the average exports with a PTA partner minus the average of the same exporter to non-PTA partners. Therefore any bilateral characteristic that is correlated with having a PTA will be reflected in that estimate. If that characteristic also affects trade independently then the estimate will be biased. So if countries that are contiguous, closer or share a language, are more likely to form PTAs and we fail to control for these variables then we overestimate the partial effect. All three of these variables can be and have been controlled for.

After controlling for basic bilateral trade determinants we are left with two potential sources of bilateral bias. First, determinants that we may consider a priori important but are currently unobservable (or poorly measured), eg, certain NTBs or other measures of policy depth. Second, bilateral controls that are potentially observable but not typically included in the earlier studies, eg, potential for conflict can be a motive for PTAs and independently affect trade.

One approach to this problem is to control for a wide range of bilateral trade determinants. This is subject to the issues we discussed under “choice of controls.” Another alternative is to use matching estimation, which generates a control group of untreated

^y We obtain a similar result if we omit the deflators in the panel and also for individual year cross sections.

Table 2 Sources of bias and heterogeneity in aggregate trade PTA estimates

	1	2	3	4	5	6				
Issue (s) addressed	MR		MR Cov. sample	MR Cov. sample OVBij/End. PTA	MR Cov. sample OVBij/End. PTA, Implem./dynamic	MR (Multilateral resistance) Cov. sample OVBij/End. PTA, Implem./dynamic, Het. PTA depth				
Approach										
Time effects	Y	Y	Y	Y	Y	Y				
Country × time effects		Y	Y	Y	Y	Y				
Bilateral effects				Y	Y	Y				
ATE _t	T	T	T	T	T + T-5 + T-10	T + T-5 + T-10				
Type PTA	All recip.	All recip.	All recip.	All recip.	All recip.	NR PTA	R PTA	FTA	CU/CM/EU	
Membership										
PTA ($\hat{\phi}$)	0.801*** (0.0421)	0.545*** (0.0408)	0.563*** (0.040)	0.369*** (0.0308)	0.60*** (0.050)	-0.006 (0.073)	0.413*** (0.068)	0.533*** 0.062	1.16*** (0.091)	
WTO ($\hat{\phi}^{WTO}$)	0.212*** (0.0273)	0.207*** (0.0622)	0.278*** (0.054)	0.086* (0.0491)	0.204*** (0.073)		0.242*** (0.073)			
Obs.	131,695	131,695	158,713	158,713	139,407	139,407				
R ²	0.63	0.73	0.71	0.82	0.83	0.83				

Data: 5-year panel 1965–2010. See the Online Appendix for data sources, summary statistics, and code availability for replication. Issues addressed: multilateral resistance (MR); covariate sample (Cov. sample); omitted bilateral variable bias (OVBij); Implementation/Dynamic effects; heterogeneous effects (Het. PTA).

Controls for specifications in columns 1 and 2: distance, border, and common language. Additional controls in column 1 GDP_i, GDP_j, and deflators. PTA in columns 1–5 corresponds to any reciprocal agreement (RPTA, FTA, or CM).

Robust standard errors clustered by bilateral pair in parentheses. Singleton observations dropped.

Columns 5 and 6 report the sum of the contemporaneous and lagged effects.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

For summary statistics, see Table A2 (<http://dx.doi.org/10.1016/bs.hescop.2016.04.013>) in the Online Appendix.

pairs that have the same predicted probability of a PTA as the treated observations. This requires a first stage probability model of PTA formation.^z A similar model is also required by the third alternative, IV estimation, where the latter also requires the standard exclusion restriction.

Both matching and IV estimates treat PTA formation as endogenous from an econometric perspective. This is also the only reasonable approach from an economic perspective but raises the question of why PTAs form in the first place. We defer detailed discussion of this point until [Section 5](#) when we will have a better understanding of various effects of PTAs. But we can discuss some potential determinants of PTAs and difficulties in addressing endogenous PTA formation.

There is currently no widely accepted model of PTA formation. Moreover, it is doubtful any such single model exists given the large diversity of agreements and policies covered in PTAs. Therefore the initial approaches to this issue have sensibly used a variety of economic and political variables as determinants of PTAs. [Magee \(2003\)](#) for example uses 2SLS in a panel with the following instruments to identify the PTA effect in the trade equation: log difference in GDP, the intraindustry trade, bilateral trade surplus, capital–labor ratio similarity, and a joint democracy dummy. These are assumed to be excludable from the trade equation but there is no test of it. The impact on the estimated trade effects of treating the PTA as endogenous is very sensitive to the exact specification leading him to conclude that “we should be cautious in using gravity equation estimates to draw strong conclusions about the effect of PTA formation on trade.”

More recent work by [Egger et al. \(2011\)](#) pursues the IV approach in a cross section of 126 countries for 2005. Relative to [Magee \(2003\)](#) they include a broader set of political similarity variables, potential bilateral costs (eg, similar language, colonial history), and control for structural multilateral terms. [Egger et al. \(2011\)](#) further assume that three indicators are excludable from the trade equation, former colony, common former colonizer, and former country (1 if a current pair of countries was once united). These factors are significant in determining the PTA and not significant in the trade equation conditional on the remaining regressors. The authors thus estimate a partial effect of PTAs between members of $\hat{\phi} = 1.15$. This is substantially larger than the 55 log points estimate they obtain if the PTA is treated as exogenous.

A final approach, used by [Baier and Bergstrand \(2007\)](#), is to explore panel data and assume that the main source of bilateral bias is time invariant. More specifically, they consider the following extension of (6), which adds a time dimension to the country effects and a bilateral fixed effect:

$$\ln T_{xmt} = \hat{\phi} PTA_{xmt} + \alpha_{xt} + \alpha_{mt} + \alpha_{xm} + \epsilon_{xmt} + u_{xmt} \quad (7)$$

^z [Egger et al. \(2008\)](#) use this approach to examine the effect of new agreements between OECD members.

Their fixed effects estimate is $\tilde{\phi}=0.46$, when only contemporaneous PTA effects are included, which is higher than the effect they obtain in the pooled sample with only time effects (0.27). This again suggests that controlling for endogeneity increases the estimated effect. However, this increase reflects both the inclusion of country-by-time effects and the bilateral effect, they have no specification that adds only the latter. To capture this we can compare columns 3 and 4 of [Table 2](#), which show that including the bilateral effect actually reduces the PTA partial effect to 0.37.^{aa}

Relative to the cross section, the approach in [\(7\)](#) has the advantage that it controls for any fixed bilateral characteristics. However, it does not eliminate the concern with potential selection bias, it simply moves it to the time-varying bilateral effect. To understand this note that the identification now relies solely on countries that entered (or left) PTAs at some point in the sample. In essence we are comparing the change in exports over a 5-year period between a pair that enters a PTA and the same countries' change in exports in that period relative to nonmembers. The question then becomes why the same country entered a PTA with m but not j in that period.

To address this concern with the timing of PTAs and whether they are correlated with factors in ϵ_{xmt} that independently raise trade we could employ IV in this panel setting. Ideally these time-varying bilateral instruments would be suggested by theory. Until this is done we cannot be completely certain about the exact interpretation of these particular estimates.^{bb}

3.2.3 Sample Issues and Heterogeneous PTA Effects

One advantage of structural gravity and its parsimonious panel specification in [\(7\)](#) is to minimize the need for covariates that may induce nonrandom sample selection. In terms of selection due to zero trade there are different approaches that have been used. This is not an issue if we are interested only on the impact of PTAs on the treated and believe that nontraders will never form a PTA. Evidence from meta-analysis indicates no significant differences between estimates that address the zeros issue.

The recent estimates described earlier address the small PTA sample issue by using a large number of agreements and focusing on estimating an average effect. This is

^{aa} Their sample comprises 96 countries in 5-year periods between 1960 and 2000 whereas ours is a non-balanced panel from 1965 to 2010 and uses a more comprehensive PTA list. For summary statistics see [Table A2](#) (<http://dx.doi.org/10.1016/bs.hescop.2016.04.013>) in the Online Appendix. The meta-analysis in [Cipollina and Salvatici \(2010\)](#) finds that estimates are on average higher in specifications that include the bilateral effect. Our sample extends to 2010 so beyond any of the studies they include.

^{bb} The approach in [\(7\)](#) would yield consistent estimates if the gain to join an agreement was log separable into a time invariant bilateral effect and country-by-time effects. For example, if Poland is a “natural” PTA partner for Germany but not for Japan then a shock that increases Poland’s gain from a PTA with *any* country in the world will generate a PTA with Germany but not Japan. This type of separability may not be satisfied in domino theories of regionalism where the bilateral component of this value can be inherently time varying. We discuss this further in [Section 5](#).

reasonable conditional on focusing on aggregate data. On the other hand we expect agreements to have heterogeneous effects, eg, certain agreement reduce tariffs reciprocally, others eliminate them altogether while others have common tariffs and/or a CM.

At one extreme there are studies that estimate PTA-specific effects and find they are heterogeneous, eg, [Eicher and Henn \(2011\)](#).^{cc} One potential downside of doing so with aggregate data is the small PTA sample bias. An alternative is to focus on average effects for broad groups of agreements. [Baier et al. \(2014\)](#) do this using the common classification in [Section 2.2](#) and find stronger effects for deeper PTAs, namely CU and FTA. We find a similar pattern when the data are extended to 2010 both for the contemporaneous effects and when we account for lagged effects (column 6 in [Table 2](#)).

In [Section 2](#), we described the variation in PTAs along policy depth and breadth. That data can be used to examine the source of heterogeneous effects. Given the large number of provisions and correlation between some of them there are two possible approaches. One is to use principal component analysis. The other is to make use of an ex ante grouping of provisions to test which are important, the taxonomy in [Section 2](#) may be useful in guiding such groupings. The data to do so are available online for the interested reader.^{dd}

Another alternative to estimate heterogeneous effects with aggregate data is to rely on theory to motivate interpretable parsimonious interactions of the PTA dummies with relevant country and/or bilateral characteristics. For example, interactions with indicators of the depth of policy such as the average level of MFN tariffs and NTBs against nonmembers.

A complementary approach to minimize the small PTA sample problem while allowing for heterogeneous effects is to use disaggregated data and specific agreements, as we will later describe.^{ee} This will be useful when we are trying to explain the magnitude and mechanism that generate the most recent PTA estimates. Before doing so we discuss the magnitude of these effects and whether they can simply be explained by tariff reductions in standard models.

3.2.4 Summary

In sum, the partial PTA effect estimates from naïve gravity approaches were heterogeneous in terms of magnitude and statistical significance, which raised the question of why countries increasingly pursue these. I argued these estimates were fragile due to different biases. Some biases were common to all gravity estimates, eg, omitting multilateral resistance, and others were specific to the topic, eg, PTA endogeneity. More recent estimates that address these and other econometric issues are more consistent in terms of their positive and significant effect and can thus better explain the fast rate of PTA formation.

^{cc} [Vicard \(2009\)](#) on the other hand does not find significant differences across different types of PTAs.

^{dd} [Egger and Wamser \(2013\)](#) find that trade is increasing in the breadth of integration agreements: from those that include only goods to services then investment and then double taxation. Not all of these are PTAs as we defined them.

^{ee} An early example that illustrates the value of doing so is [Clausing \(2001\)](#), which we discuss in [Section 5](#).

I highlight a few key issues from the prior discussion. First, any gravity estimate must start from a well-defined theoretical model such that the identification assumptions and interpretation of the PTA effect are clear. Second, aggregate gravity estimates should focus on average effects rather than individual PTA effects—the latter require more detailed data. Third, the panel results in [Table 2](#) indicate that controlling for multilateral resistance and bilateral fixed effects reduces the magnitude of the PTA effect. Moreover, as we now discuss, the recent evidence points to dynamic trade effects of PTAs that are stronger for deeper agreements.

3.3 A PTA Trade Elasticity Puzzle?

The theory consistent estimates of the average PTA effect that address the key econometric issues we discussed are robustly positive. If PTAs reduce bilateral trade costs then this positive partial effect is not surprising and rules out few, if any, PTA theories. For example, a positive effect is consistent with the standard static view of PTAs as a reduction in a preferential tariff. However, the standard static models are unable to explain other important pieces of evidence. First, in [Section 2](#) we described the recent emphasis on negotiating policies other than tariffs. Second, the finding above that CU and CM have substantially stronger trade effects indicates other policies may play a role.

In this section, we discuss the plausible range of magnitudes of long-run PTA effects from recent estimates and argue that accounting for them in a standard static model would require either an implausibly large (i) trade elasticity or (ii) reduction in preferential tariffs. We then provide direct evidence that neither of these two conditions are present in this data and that the PTA effect remains even after tariff reductions are taken into account.

3.3.1 *The Magnitude and Timing of Aggregate PTA Effects*

PTAs are highly persistent and thus it is important to understand their long-run impact. The fact that cross-sectional estimates are larger than the panel ones indicates that the trade effect of PTAs may be larger in the long run. Two plausible motives would be that policies are phased-in over time and/or there may be dynamic effects that build up gradually.

[Baier and Bergstrand \(2007\)](#) capture the long-run effects by reestimating (7) including not just a contemporaneous PTA effect but also an indicator if the PTA is present for at least 5 years and at least 10 years. They find the PTA effect doubles after 5 years (from 0.28 to 0.55) and is triple the short-run effect after 10 or more years. Their long-run estimate is 0.76, which is higher than the average effect when lagged terms are omitted.^{ff}

^{ff} This difference may be due to the substantial number of new PTAs in the latter part of the sample such that the average estimate without lagged terms reflects their short-run effect. The lagged effect could also reflect positive correlation in the error term, which in some cases can be ameliorated by first differencing, but applying the latter yields similar estimates.

In column 5 of Table 2 we include a similar lag structure and find a similarly large increase in the average PTA effect in the data for 1965–2010. A similar increase is present for each of the separate types of reciprocal PTAs (column 6). Thus panel analysis with only a contemporaneous PTA effect may generate a downward biased estimate of the long-run effect, particularly in short panels.^{gg}

In sum, recent estimates of the average PTA effect that are consistent with the structural gravity requirements and take long-run effects into account are fairly high. This is true for panels, with both Baier and Bergstrand (2007) (table 5) and Anderson and Yotov (2016) finding a value of around 0.76, and even higher in a cross section, 1.15 (Egger et al., 2011, table 2).

3.3.2 Puzzle: Definition and Existing Evidence

Given the high estimates just described we now ask two related questions. First, at a given trade elasticity, what is the maximum that observed tariffs can explain of the PTA effect? Second, at the current observed MFN tariffs, how high would the trade elasticity have to be in order to fully explain that effect? We find that observed tariffs can only explain a fraction of the PTA effect if we use a standard value of the trade elasticity and that to fully explain the PTA effect requires an elasticity considerably higher than standard estimates, hence the elasticity puzzle.

Traditional theories of PTAs focus on its role in reducing tariff barriers between members. If that is all a PTA did then, given a log change $\widehat{\tau}_{xm}$ in the ad valorem tariff factor, we would obtain $\widehat{T}_{xm} = \widehat{\phi}_{xm} = -\varepsilon \widehat{\tau}_{xm}$, where $\varepsilon > 0$ is the absolute value of the variable trade cost elasticity. We say there is an elasticity puzzle if the estimated effect, $\widehat{\phi}_{xm}$, can only be reconciled with the prediction from such a model by using an implausible trade cost elasticity.

In standard models that yield a gravity equation this elasticity is a parameter. In a single sector Armington model it is proportional to the constant elasticity of substitution across varieties, $\varepsilon = \sigma - 1$. In a Melitz–Chaney framework $\varepsilon = k$, the Pareto productivity dispersion parameter.^{hh} The estimated ad valorem trade cost reduction of PTAs is then defined as

$$\widetilde{PTA}^{ave} \equiv \widetilde{\phi} / \varepsilon \quad (8)$$

The trade elasticity plays a key role in recent quantifications of welfare effects of trade and there is a range of estimates for it. Most estimates fall in the 3–7 range and for this

^{gg} This is clearly reflected in the higher value of the full panel estimate in column 5, 0.6, with either of the corresponding coefficient estimates for alternative subperiods, eg, 1990–2010 and 1965–85, which are presented in columns 2 and 3 of Table 3. Thus the dynamic effects introduce an additional potential cost of any sample selection that shortens the length of the panel, eg, due to missing data.

^{hh} This mapping assumes that tariffs are imposed as export costs. If we impose them on the consumers then the formulas must be adjusted, for example $\varepsilon = k\sigma/(\sigma - 1)$.

discussion we follow Costinot and Rodriguez-Clare (2014) and Head and Mayer (2014) and focus on an intermediate value, $\varepsilon = 5$. At this elasticity we obtain $\widetilde{PTA}^{ave} = 1.15/5$. That is the partial trade effect from a PTA is equivalent to eliminating a 26% ad valorem tariff ($\exp(1.15/5) - 1$). For the panel estimate it is 16%.

Kee et al. (2009) calculate the trade restrictiveness of MFN tariffs for each country, the average of this measure across countries in the world in 2009 was about 7.4%.ⁱⁱ So even full elimination of MFN tariffs between PTA members is unable to account for its partial trade effect.^{jj}

In practice, how large are the actual changes in tariff barriers between members after a PTA? The surprisingly little systematic research on this point shows that it is considerably less than what would be implied by a complete removal of tariffs. The WTO's (2011) report finds that a large share of trade in PTAs occurs in tariff lines that already have zero MFN tariffs and that a number of products where the MFN tariff is high are excluded from PTAs. Overall it calculates a preferential margin of 2.1 percentage points for tariffs. This margin is higher for countries with higher MFN tariffs such as Mexico (9.3), it is 4.9 within the European Union and only 0.7 for the United States. Recent work by Hayakawa and Kimura (2014) finds that on average PTAs lead to reductions of 2% for tariffs. So these tariff reductions can only explain a small fraction of \widetilde{PTA}^{ave} .

What alternative values of the trade elasticity are required in order to account for the PTA effect? Under a full removal of the average observed MFN tariffs, 7.4%, the required elasticity is between 10 for the panel estimate and 15 for the cross section. The required elasticity is substantially higher if we employ the measured preferential tariff liberalization of 2%.

3.3.3 Puzzle: Novel Evidence

We now ask if the puzzle persists in a more recent sample. More importantly, we use tariff data to (i) estimate the relevant trade elasticity and (ii) directly control for this channel of PTAs. This allows us to answer how much larger the tariff reduction and/or trade elasticity have to be to explain the PTA effect while using a unified dataset and methodology instead of piecing together estimates from different studies.

In column 1 of Table 3 we replicate column 5 of Table 2 for comparison, which estimates the long-run effects of the RPTA variable for 1965–2010. The tariff data are available after 1988 so the second column reestimates over the 1990–2010 period. We see a pattern similar to the full sample but with smaller effects.^{kk} The estimate for the subsample

ⁱⁱ Using the July 2012 version of the data available at <resources.worldbank.org/INTRES/Resources/469232-1107449512766/OTRI2009.xlsx >

^{jj} Head and Mayer (2014, p. 165) reach the opposite conclusion because they use a substantially lower $\tilde{\phi} = 0.28$, which is the median from their meta-analysis that includes studies subject to the biases previously discussed.

^{kk} This reduction in the average effect is also present if we use the subsample 1960–85, in column 3, and suggests that the shorter panels are unable to capture longer run effects.

Table 3 PTA average effects on trade and applied tariffs

	1	2	3	4	5	6	7	8	9	
Dependent variable	ln imports					ln applied tariff				
Sample	1965–2010	1990–2010	1965–85	1990–2010, tariff sample						1990–2010
PTA LR	0.60*** (0.050)	0.427*** (0.0623)	0.382*** (0.101)	0.264*** (0.0658)	0.210*** (0.0651)			−0.0174*** (0.00269)		
WTO LR	0.204*** (0.073)	0.146 (0.103)	−0.252 (0.169)	0.378** (0.191)	0.380** (0.190)	0.378** (0.191)	0.380** (0.190)	0.0008 (0.0066)	0.0007 (0.0066)	
ln applied tariff					−3.124*** (0.201)		−3.113*** (0.202)			
Nonreciprocal PTA LR						0.0887 (0.0641)	0.0486 (0.0642)		−0.0129*** (0.00227)	
Reciprocal PTA LR						0.196** (0.0971)	0.174* (0.0959)		−0.00717* (0.00391)	
FTA LR						0.211*** (0.0745)	0.133* (0.0737)		−0.0250*** (0.00311)	
Customs Union Common mkt, currency LR						0.733*** (0.120)	0.609*** (0.120)		−0.0401*** (0.00516)	
Obs.	139,407	96,458	41,565	64,270	64,270	64,270	64,270	64,270	64,270	
R ²	0.822	0.843	0.836	0.871	0.872	0.871	0.872	0.674	0.675	

Data: 5-year panel 1965–2010. See the Online Appendix for data sources, summary statistics, and code availability for replication.

All specifications include importer-year, exporter-year, and bilateral fixed effects.

Robust standard errors clustered by bilateral pair in parentheses. Singleton observations dropped.

ln applied tariff defined as the $\ln(1 + \tau_{xmt})$ where τ_{xmt} is the average bilateral applied tariff at t , which captures any preferential tariffs.

WTO LR and all PTA LR variables refer to sum of coefficients at T, T-5, T-10.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

For summary statistics, see Table A3 (<http://dx.doi.org/10.1016/bs.hescop.2016.04.013>) in the Online Appendix.

of countries with available tariff data is in column 4 and implies a PTA effect of 0.26. Our objective is to determine if this PTA effect, which is smaller than the ones using the longer panels, can be explained by preferential tariff reductions.

Not all PTAs eliminate tariffs fully; and even FTAs that aim to do so include provisions, such as rules of origin, that must be satisfied for a good to be eligible for the preference. To capture the change in tariffs due to an agreement we must therefore control for the effectively applied tariff faced by an exporter x , which is the average duty it paid when selling to m at time t , τ_{xmt} . This measure varies bilaterally and over time and thus its effect, the trade elasticity, can be identified using the augmented version of (7) shown later. This specification includes lagged PTA effects and controls for changes in tariff protection, so the remaining PTA effect, denoted by $\hat{\phi}_{-\tau,i}$, reflects everything except applied tariff barriers.

$$\ln T_{xmt} = \sum_{i=0,5,10} \hat{\phi}_{-\tau,i} PTA_{xmt-i} + \epsilon \tau_{xmt} + \alpha_{xt} + \alpha_{mt} + \alpha_{xm} + \epsilon_{xmt} + u_{xmt} \quad (9)$$

In column 5 we find that, after controlling for tariffs, PTAs continue to increase bilateral trade significantly after at least 10 years. The effect is about 5.5 log points lower than the estimate in column 4, over the same sample but without tariffs. This difference is consistent with our estimated tariff elasticity, $\tilde{\epsilon} = 3.1$, provided the preferential tariff reduction is around 1.7 log points.^{ll}

We can verify the reduction in preferential tariffs generated by PTAs directly in our data by using a specification similar to (9) with τ_{xmt} as a dependent variable. In column 8 we do so and find an average preferential reduction of almost 1.7 log points.^{mmm} In sum, given our estimated elasticity, the tariff reduction would need to have been about five times larger than observed to account for the PTA effect in this sample. Alternatively, given the observed preferential tariff reductions, the trade elasticity would have had to be about five times larger than what we estimate it to be, ie, over 15, which is considerably higher than the standard values.

In column 6, we disaggregate the RPTA variable into the three common categories previously described and find a positive and significant effect for each in the 1990–2010 subsample; controlling for the tariff in column 7 reduces their magnitude but does not change their sign or standard error. The tariff reduction is largest for CU but it still only explains less than one-fifth of the CU partial trade effect. The average tariff change in FTAs explains around a third of the effect. Note also that even though NRPTA

^{ll} There is a potential endogeneity concern arising from the tariff. But, similarly to the PTA, this concern is at least partially addressed by two controls. First, the importer-time effects, which control for any aggregate reform. Second, the bilateral fixed effect, which among other things controls for the composition of bilateral trade and resulting differences in average tariffs between countries that eventually form a PTA.

^{mmm} This allows for phase-in effects. Hayakawa and Kimura (2014) find a similar effect using a yearly sample starting in 1995. Their specification does not control for importer-year and exporter-year effects as we do.

experience a significant tariff decrease there is no significant PTA effect, which may be due to the uncertain nature of those tariffs.

In sum, the PTA effect is not fully explained by preferential tariff reductions. This justifies the widespread use of dummies in the gravity approach to capture other channels through which PTAs can increase bilateral trade, but it also begs the question of what those channels are. Moreover, we estimate that recent PTAs have very modest effects on tariff reductions, which reflects the fact that MFN tariffs are already low, as noted in Section 1.

3.3.4 Features of Possible Explanations for the Elasticity Puzzle

The last results suggest that either PTAs increase the trade elasticity with respect to tariffs and/or they reduce trade costs beyond tariffs. Our objective here is not to list explanations and resolve the puzzle but to highlight important features of potential explanations. In Sections 4 and 5, we analyze the deeper and broader economic cooperation in modern PTAs and will point out which aspects of a richer economic and/or policy structure can contribute to explain the elasticity puzzle.

Recall that we defined the elasticity puzzle by comparing the magnitude of the estimated effect, $\hat{\phi}_{xm}$, and the prediction arising from a class of models. What are the characteristics of the models where there is an elasticity puzzle? First, these models generate a structural gravity equation, so we can estimate the partial PTA effect using $\hat{\phi}_{xm}$ as described in (7). Second, they assume a particular economic and policy structure such that this effect is linear in the trade elasticity and tariff change:

$$\hat{\phi}(\pi_{xm}^{PTA}) = -\varepsilon \hat{\tau}_{xm}. \quad (10)$$

More specifically, this structure requires (i) constant trade elasticity (over goods, policies, and time), (ii) trade only in final goods (otherwise $\hat{\tau}_{xm}$ would have to reflect some weighted average), (iii) deterministic policy, and (iv) tariffs as the only instrument affected by PTAs.

Thus in order to explain the puzzle we should consider models that still generate a gravity structure, such that the estimate of $\hat{\phi}_{xm}$ as an average effect is still valid and feature any combination of (a) additional reductions in policy frictions and/or (b) a tariff elasticity function, $\varepsilon(\pi_{xm}^{PTA})$, that is higher than the average in the literature.^{nm}

In subsequent sections, we examine how important each of the assumptions listed in (i)–(iv) are and what may be relevant additional frictions and sources of higher trade

^{nm} The NTBs can also include measures that affect fixed costs. Thus we could alternatively ask what the percent reduction in those costs, \hat{f}_{xm} would have to be to explain a given PTA effect. This requires a specific model and elasticity, in a Melitz–Chaney framework the relevant elasticity is $\frac{\varepsilon}{\sigma-1} - 1$. After tariffs the unexplained effect is $\tilde{\phi} - \varepsilon \hat{\tau}_{xm} = \left(\frac{\varepsilon}{\sigma-1} - 1\right) \hat{f}_{xm}$.

elasticities. Briefly, we may obtain higher trade elasticities if we relax the economic structure constraints (i) and (ii) by allowing for trade in intermediates and certain export investments (Section 4) or FDI (Section 5), for example. In Section 4, we also relax (iii) and (iv) and allow for additional policy frictions in the form of observable NTBs and trade policy uncertainty (TPU). We also show that policy uncertainty combined with export investments can generate higher elasticities, which along with lower expected protection, implies larger trade effects of PTAs.

3.4 Ex Post Estimates of WTO Trade Effects

In the introduction and in Section 2, we noted that a substantial fraction of PTAs are between WTO members. Thus to identify the partial effect of PTAs it is important to control for participation in the WTO, which has not always been the case in the literature but is what we do in Tables 2 and 3. We briefly discuss the partial trade effect of GATT/WTO membership, which is interesting in of itself and as a reference point for the effects of PTAs. The basic approach is the one we outlined for PTAs. To distinguish between estimates we refer to the partial effects for the WTO as $\hat{\phi}^{wto}$ and continue to use $\hat{\phi}$ for the other agreements.

Rose (2004) first examined the WTO trade effect. His conclusion was that joining or belonging to the GATT/WTO did not have a significant impact on bilateral trade. His baseline approach falls under the naïve gravity group, he uses total trade, focuses on pooled data with time effects and country controls such as GDP. In the one specification with country fixed effects the results are actually positive and significant but modest ($\hat{\phi}^{wto} = 0.15$). But these are not time varying and so do not fully account for the multilateral terms. In another specification Rose controls for country pair effects (Table 3), again the estimates are higher on average, particularly for earlier rounds (from start of GATT up to Kennedy round the effect is 0.24–0.76) but because they are separately estimated for each round, there is variation across the estimates and at least one is negative.

Subramanian and Wei (2007) also use a gravity approach but conclude that the GATT/WTO promoted trade strongly, by about \$8 trillion in 2000 alone, and unevenly. I highlight the following factors contributing to the difference in the results relative to Rose (2004). The authors' use of (i) country-by-year effects to control for the structural multilateral terms and (ii) imports instead of trade to capture potential asymmetries in the effects of developed and developing countries. This is important because developing countries have traditionally liberalized less during trade rounds and the authors find that the WTO effect is in fact insignificant for them but positive and significant for developed countries. Another difference is these authors' different definition of WTO and PTA membership whereby if a country pair is a member of the WTO and a PTA then they code the former as 0 and the latter as 1, ie, they are mutually exclusive. Their argument for doing so is that it identifies the "pure" WTO effect without

confounding it with that of potential future PTAs. Under this definition they find a large average PTA effect and also a WTO effect for developed countries.

Eicher and Henn (2011) point out an issue that has important implications for the interpretation of the earlier estimates. They show that if we include a separate developed country PTA effect then the Subramanian and Wei (2007) classification for this and developed WTO member is perfectly collinear with the country–year effects. This arises because developed countries first enter the WTO and only then join PTAs with each other, thus they argue the developed country WTO effect is really a PTA effect. They then reclassify the variables and estimate the following effects $\{(PTA, \text{non-WTO}) = 0.7, (PTA, \text{WTO}) = 0.52, \text{WTO} = -0.07\}$. Thus PTAs have positive and significant effects for WTO and non-WTO pairs alike but conditional on them there is no additional average WTO effect. They find heterogeneity in the WTO effect but not along the developing, developed margin. For example, there are modest effects of WTO accession for countries that had higher imports upon accession and thus higher potential market power and initial tariffs.

We use the data for the WTO in Eicher and Henn (2011) extended to 2010 and also allow for membership in the WTO and PTAs not to be mutually exclusive. In Table 2, we can see how various econometric concerns affect $\tilde{\phi}^{wto}$: controlling for bilateral effects (column 4) substantially reduces the effect, but it remains positive and significant. It is also worth noting that, similarly to PTAs, the WTO effect is significantly larger after at least 10 years of membership, $\tilde{\phi}^{wto} = 0.2$. This effect increases further if we disaggregate the PTAs by depth and control for nonreciprocal preferences, such as GSP, as we see in column 6. If we were to ignore dynamic effects and restrict the sample to end in 2000, as in Eicher and Henn (2011), then we also find an insignificant WTO effect. This points to the importance of accessions under the WTO period, which required additional commitments, and possibly the dynamic effects of the full implementation of the UR by the late 1990s.

One possible reason for the difficulty in identifying robust average WTO effects is their heterogeneity. For example, upon accession some countries may not have liberalized, or may have liberalized unilaterally. In Table 3, column 8, we find that the average tariff in 1990–2010 does not change significantly between new pairs of WTO members. Subramanian and Wei (2007) already provided some evidence for heterogeneous effects. Chang and Lee (2011) go one step forward and show that interacting the WTO membership with a variety of covariates (income, geography, etc.) yields significant effects.^{oo} Dutt et al. (2013) also find that the WTO effects are heterogeneous along the intensive margin (negative) and extensive margin (positive).

^{oo} Thus they argue for a more flexible nonparametric approach that matches treated pairs and compares their mean trade with untreated ones. In a sample similar to Rose (2004) they find stronger and more robust average WTO effects.

Another difficulty in identifying an average WTO effect is that the standard estimation approach is based on accession and thus excludes the original GATT signatories, ie, several large industrialized countries. Pre-GATT data would circumvent this issue. An alternative avenue is to explore disaggregated data and estimate the impact of specific policies such as tariff bindings, as discussed in [Section 4](#).

3.5 General Equilibrium Trade and Welfare Effects of PTAs

Our focus on the partial trade effects is driven by two factors. First, it has been the focus of much of the research. Second, the recent estimates are robust to alternative modeling assumptions. However, given the prevalence of PTAs and the large partial effects found it is important to examine their general equilibrium effects. Early work using CGE models did precisely this with mixed success (cf. [Hertel et al., 2007](#)). Here we focus on recent approaches using new quantitative trade models.

We are interested in \hat{T}_{xm}^{ge} in (2), which requires us to go beyond the partial effect and compute the effects on X_x and M_m . While different models generate the same structural gravity in (1) they differ in their implications for the exporter and importer terms, and thus yield different GE estimates. The basic approach is to take the estimated change due to a PTA and translate it into an ad valorem equivalent using a particular trade elasticity and then ask what the counterfactual values of X_x and M_m would be. In practice this entails using the structure of the chosen model to calculate the impact of this cost change on income, production, and the multilateral resistance terms defined in [Section 3.2](#) to derive new trade flows.

[Egger et al. \(2011\)](#) estimate a partial effect and use it to compute the counterfactual trade outcomes without PTAs in the context of an Armington model. The difference between the counterfactual and observed trade yields an average increase in trade of 102% between PTA members. While the partial effect is estimated to be the same, the general equilibrium effect, \hat{T}_{xm}^{ge} , is quite heterogeneous because countries have different characteristics such as size and openness, the standard deviation is 95, but nearly all pairs have positive effects. Given this heterogeneity, it would have been useful to calculate the aggregate trade effect of PTAs, since some of the larger percent increases could apply to pairs with small amounts of trade.

It is also important to note that this approach allows us to quantify the impact of PTA on trade with nonmembers, which falls by 9% on average with declines for about two-thirds of the pairs. Thus PTAs not only create trade between members but also divert some away from nonmembers. We return to the issues of trade creation vs diversion and third country externalities in [Sections 5 and 6](#).

This analysis can in principle be extended to analyze other counterfactuals, eg, the impacts of removing specific PTAs. Moreover, the class of model estimated has a simple measure of welfare change that depends only on the change of trade openness, which can

be computed from the counterfactual, and the partial trade elasticity. It would be interesting to calculate these and examine if particular PTAs generate higher welfare for a country than others. For example, some models predict higher welfare gains from regional PTAs, which may be less prone to trade diversion if most of the trade is regional. [Krishna \(2003\)](#) examines this question using a different approach. He estimates a demand system resulting from a general equilibrium perfect competition model to obtain the price effects of a US preferential tariff reduction and does not find evidence that the welfare gains are larger if the PTAs are with closer countries.

[Caliendo and Parro \(2015\)](#) evaluate the impact of a specific agreement, NAFTA, using a multisector Ricardian model that incorporates sectorial linkages and intermediates goods' trade. They also find very heterogeneous trade effects for members ranging from 118% for Mexico to 11% for Canada. These translate into modest welfare increases for Mexico, 1.3%, and small changes for the United States and Canada. As [Costinot and Rodriguez-Clare \(2014\)](#) notice the quantitative trade predictions are better aligned with the data than the predictions of earlier CGE models, but the reason for this is unclear.^{PP} Further quantitative work will likely bridge some of the gap between the parsimony and clear microfoundations of these new quantitative approaches and the level of detail in typical CGE approaches.⁹⁹

[Caliendo et al. \(2015\)](#) use a framework similar to [Caliendo and Parro \(2015\)](#) to quantify the trade welfare gains of the multilateral liberalization in the Uruguay Round and PTAs in 1990–2010. They find the welfare effect of PTAs was very small, 0.3%, particularly when compared to the almost 6% increase found for the UR. Underlying the small welfare gains for PTAs is the small trade effects of these agreements in their model. How can these small trade effects consistent with the earlier evidence we provided? I believe this is due to their measurement of PTAs as simply a tariff reduction. As our evidence in [Table 3](#) shows, PTAs had small effects on applied tariffs in 1990–2010 and when these small changes are applied to a trade elasticity of 5.5 (their measure for manufacturing) the predicted PTA effect in their model is very modest.^{TT}

[Anderson and Yotov \(2016\)](#) provide comprehensive estimates of the GE effects of all PTAs in 1990–2002. They find large trade and welfare effects of PTAs in contrast to [Caliendo et al. \(2015\)](#). The difference in the results is at least in part driven by their distinct approaches in capturing the magnitude of the PTA shock. Anderson and Yotov

^{PP} [Kehoe \(2005\)](#) argues that the earlier CGE models can only deliver the observed aggregate effects of NAFTA if we assume that trade elasticities are unreasonably high and in the “wrong” sector.

⁹⁹ [Levchenko and Zhang \(2012\)](#) use a related approach to compute the welfare gains of Eastern European integration with the European Union. They find Eastern European country welfare increases over 9% on average, with the largest gains going to those with comparative advantage and technology most different from countries in Western Europe, which are barely affected (0.16% change in welfare).

^{TT} See also [Spearot \(2016\)](#) who quantifies the effect of multilateral liberalization and uses the structural model to evaluate the prospective effects of the Trans-Pacific Partnership.

retrieve the PTA average partial effect from a structural gravity estimate rather than using only the part that is implied by observed tariff changes. This reinforces the importance of expanding our view of PTAs beyond changes in applied tariffs.

4. ECONOMIC EFFECTS OF DEEPER PTAs

To understand the effects of recent deeper PTAs we must augment the policy and economic structure relative to the traditional view. Here I focus on augmenting policy to include current NTBs and uncertainty about future policies. The economic structure is augmented to consider investments in export activities and intermediate goods. I highlight how these features contribute to understanding the magnitude of the estimated trade effects and their heterogeneity across types of PTAs. Moreover, I argue that to better understand and quantify the impacts of deeper PTAs we must model and estimate the effects of specific policies at a disaggregated level and describe how this has been done in a specific context: TPU.

4.1 Nontariff Barriers

Section 2 describes how a large fraction of PTAs also aim to reduce NTBs. The NTB policy vector, ν , includes barriers that are currently applied, eg, costly customs procedures, import licenses, and harmonization of product standards, as well as rules about contingent protection, eg, antidumping and countervailing measures. These are diverse in terms of their mechanism and different subsets of them are included in different PTAs; so any definitive answer on the motives for their inclusion and their impact on trade requires detailed studies of specific PTAs. Here we ask two narrower questions.⁸⁸ First, how much do PTAs reduce current NTBs? Second, how can the aggregate trade effects of PTAs via NTBs be estimated and to what extent can they account for the elasticity puzzle?

If we had a time-varying ad valorem equivalent measure of bilateral NTBs, ν_{xmt}^{ave} , then we could apply an approach similar to the one used to determine the effect of PTAs on tariffs in Table 3. There are three important measurement issues in obtaining such a measure. First, information about many NTBs are recorded simply as binary variables. Second, it is not obvious how to aggregate the effects across disparate NTBs and goods. Third, some NTBs are not measured and/or do not vary bilaterally and others may not even be recorded.

Kee et al. (2009) address the first two measurement issues. They regress aggregate imports for each country on a dummy variable equal to unity in the presence of a core NTB. They do this separately for each good to obtain the trade effect of core NTBs and then divide it by an estimated import demand elasticity to obtain an ad valorem

⁸⁸ Ederington and Ruta (2016) provide a comprehensive analysis of nontariff measures.

equivalent.^{tt} These estimates can be consistently aggregated to compute the uniform tariff ad valorem equivalent that would generate the same aggregate imports. They combine this and MFN tariffs to calculate their aggregate impact on trade, which is about 12% on average in 2009, whereas the corresponding number for the MFN tariff alone was 7.4%. Thus on average if PTAs eliminated all applied tariff and core NTBs then their ad valorem equivalent reduction would be at most twice as high as if they eliminated only tariffs.

An alternative approach is to further explore the predictions of models to infer *all* applied ad valorem bilateral trade costs. [Novy \(2013\)](#) shows that different trade models that generate a structural gravity equation imply that changes in a country's trade with a partner relative to its domestic trade will reflect changes in bilateral trade costs. These costs can then be translated into an ad valorem equivalent using a particular trade elasticity. After netting out the observed bilateral tariffs we obtain an ad valorem equivalent of all nontariff bilateral trade costs. These reflect both the NTBs we are interested in as well as any behind-the-border policies, changes in transport infrastructure, information costs, etc. More specifically, when there is a constant trade elasticity we have $\phi_{xm} = t_{xm}^{-\varepsilon}$ where t_{xm} is the unobserved total ad valorem bilateral cost. We can then use the structural gravity Eq. (1) and trade data to compute the following relative geometric average of all bilateral trade costs:^{uu}

$$\bar{t}_{xm} \equiv \left(\frac{t_{xm} t_{mx}}{t_{xx} t_{mm}} \right)^{1/2} = \left[\frac{T_{xm} T_{mx}}{T_{xx} T_{mm}} \right]^{-1/2\varepsilon}$$

[Novy \(2013\)](#) assumes $\varepsilon = 7$ and calculates \bar{t}_{xmt} for a panel of countries over time.

[Hayakawa and Kimura \(2014\)](#) use $\ln \frac{\bar{t}_{xmt}}{(\tau_{xmt} \tau_{mxt})^{1/2}}$ as the measure of nontariff bilateral

costs for manufacturing products and find it is 2.1 log points lower for countries that enter a PTA, slightly more for WTO members.^{vv} The corresponding effect for tariff reductions due to a PTA was 2 log points. They do not examine the trade effect of PTAs in their sample or estimate the trade elasticity directly. So at most we can say that at the elasticity used, 7, the estimates imply PTAs increase trade by 28 log points, roughly half of it due to tariff reductions. The other half is accounted by reductions in *all* applied bilateral trade costs, of which NTBs as we defined them earlier are only a fraction.

^{tt} They include price control and monopolistic measures, technical regulations, and quantity restrictions.

^{uu} We can see this using the standard gravity framework [Section 3](#). Solving Eq. (1) for the bilateral market access function between two countries and between themselves we obtain $\frac{\phi_{xm} \phi_{mx}}{\phi_{xx} \phi_{mm}} = \frac{T_{xm} T_{mx}}{T_{xx} T_{mm}}$. Thus, using observed trade flows between and within countries we can compute the geometric average of bilateral access between a pair x, m relative to their internal market access.

^{vv} The result is obtained by regressing the measure on lagged PTA indicators, a bilateral fixed effect and time effects for a sample of up to 158 countries yearly from 1995 to 2010.

The evidence above suggests that incorporating applied NTBs can contribute to but not fully account for the trade impacts of PTAs described in [Section 3](#). Recall that our estimates in [Table 3](#) required a preferential reduction in protection at least five times higher than observed. But incorporating the ad valorem equivalent of NTBs leads to at most a doubling of protection. This is true whether we define NTBs narrowly, as in [Kee et al. \(2009\)](#) or broadly, as in [Hayakawa and Kimura \(2014\)](#).

The NTB-related effects of PTAs have attracted particular interest in the context of TTIP. [Francois et al. \(2013\)](#) estimate this agreement can increase European Union exports to the United States by 28% and increase its GDP by 0.5% with up to 80% of this arising from a reduction of 25% in NTBs. Ex ante estimates of such trade effects require us to specify at a minimum (i) a particular model, (ii) a channel through which a specific NTB operates, and (iii) a trade elasticity and associated ad valorem change. [Francois et al. \(2013\)](#) use a version of GTAP with a rich economic structure and assume the 25% NTB reduction takes the form of a marginal cost reduction (implemented as an iceberg trade cost or a tax). But it is also reasonable that NTBs take the form of a fixed cost in which case the implied trade elasticity would be different. The magnitude of the NTB reduction is guided by firm survey data on perceived costs of exporting to different markets, which are combined with gravity estimation to generate an ad valorem equivalent.

A useful approach to measuring the impact of NTBs is to focus on a subset and examine disaggregated data as done by [Chen and Mattoo \(2008\)](#). They explore industry data for a panel of 42 countries over 1986 to 2001 and find that agreements including harmonization and mutual recognition of product standards increase bilateral trade between members.^{ww} Additional work along these lines but that would simultaneously include tariff barriers to estimate trade elasticities can generate credible ad valorem equivalent measures of changes in NTBs.

In sum, incorporating NTBs explicitly in PTA models can contribute to explain their trade impacts, even if only partially. Doing so requires careful modeling and measurement of the type of barrier and the channel(s) through which it affects trade. This is fertile ground for future work. Reducing NTBs can be particularly important when firms rely heavily on intermediates and/or can rearrange their production structure across borders, an issue to which we now turn.

4.2 Intermediates and Vertical Integration

A substantial fraction of trade takes the form of intermediate goods. Moreover, a frequently cited reason for PTAs is to allow members to break up production to either

^{ww} They find it may divert it from nonmembers if there are strict rules of origin. The effect of NTBs on nonmembers is an important area for research as we discuss in [Section 6](#).

explore economies of scale or reduce production costs by setting up labor intensive activities in lower wage countries (WTO, 2011).^{xx}

The impact of a tariff reduction on intermediate goods can easily be magnified when they crisscross borders at different stages of production. Yi (2003) shows how vertical integration can magnify the impacts of tariffs, particularly at low protection levels (when there is already a reasonable amount of specialization). To my knowledge there is no direct empirical test of this channel in the context of PTAs. Doing so requires us to consider an economic structure that augments the traditional models used to evaluate PTAs to incorporate intermediates and allow for nonlinear trade cost elasticities with respect to trade costs such as tariffs. I conjecture that two basic predictions would result from such a model and support for them would be consistent with the magnification hypothesis. The first prediction is that trade elasticities are decreasing in trade costs such as tariffs and the second one is that PTAs would increase trade by more in industries where intermediates are more important.

Using the aggregate data from Section 3, we find some support for the first prediction. More specifically, we rerun the specification in (9) including a quadratic tariff term and

obtain the following marginal effect: $\frac{\partial \ln T_{xm}}{\partial \tau} = -5.8 + 11\tau$, which is precisely estimated.

This implies the elasticity at zero tariffs is greater than its value at the sample median tariff of 4.5 log points, ie, $\varepsilon(\tau=0) = 5.8 > 5.3 = \varepsilon(\tau=4.5)$. At the 75th percentile of tariffs the elasticity is 4.5. The estimated PTA indicator coefficient is now smaller, indicating that it was accounting for some of this nonlinear tariff effect. This may be consistent with the intermediates magnification channel, but it is also consistent with PTAs increasing trade elasticities due to a different channel, eg, by reducing uncertainty about future tariff changes.

Orefice and Rocha (2014) examine the trade effect of 66 PTAs via a gravity equation between 1980 and 2007 estimated separately for final goods and intermediates (ie, parts and components). They find nearly the exact same effect for each type of industry. They also test and find that deeper PTAs, which address issues that could promote the integration of production, do *not* have a stronger effect on intermediates than on final goods.^{yy} Moreover, those effects are similar in more recent subsamples. While the effects of deeper PTAs are similar across the types of trade, the authors find that the probability of their formation is higher between countries with a larger share of trade in parts and components, particularly if one country is low income. Thus the potential to share production networks may influence the selection of PTA partners.

^{xx} Blanchard et al. (2016) provide evidence that supply chain linkages affect trade policy, as discussed in Section 5.

^{yy} The PTA depth measure for 66 agreements is based on the principal components of the WTO+ and WTO-X categories described in Section 2.

The presence of intermediates that can be sourced internationally complicates the measurement of the effects of PTAs. For example, the magnification hypothesis arises in part because trade flows are measured in terms of gross values. The magnification effect of a tariff reduction would not necessarily be present if we measured the value added in production by the PTA partner. [Johnson and Noguera \(2014\)](#) construct bilateral measures of value added trade for 42 countries and document its evolution between 1970 and 2009 and across countries.

They find an average PTA effect of 0.5 for gross exports, similar in approach and magnitude to what we report in column 5 of [Table 2](#). When they instead use value added exports as the dependent variable the average PTA effect is 0.39. The difference between gross and value added exports is statistically significant and indicates that PTAs lead to an increase in the share of trade that embodies intermediates sourced from another country. This could be because country x is now sourcing more intermediates from m and then selling the final good to m . But it is also consistent with a third country setting up new production in x to assemble the good and use x as an export platform to m . We would need additional information about the full supply chain of the product to distinguish between these alternatives and determine if the evidence supports the magnification hypothesis. [Johnson and Noguera \(2014\)](#) also provide evidence that in the context of their model PTAs reduce ad valorem bilateral trade costs and tend to do so by more for trade in intermediates, 20–25%, than in final goods, 12–21% (after 15 years).^{zz}

In order to understand the economic mechanism and role of specific policy changes in PTAs it is useful to consider specific agreements. This permits a more detailed production and policy structure where key parameters such as trade elasticities can be consistently estimated and then employed for quantification. A recent example of such an approach is [Caliendo and Parro \(2015\)](#). They report that in 1993 between 72% and 82% of imports of NAFTA countries from each other took the form of intermediate goods. They build a multisector model with intermediates and estimate heterogeneous trade elasticities across sectors. They find that incorporating intermediates increases the aggregate trade effects of NAFTA's tariff reductions. It would be interesting to isolate the importance of intermediates in other agreements and incorporate changes in NTBs as well.

4.3 Trade Policy Uncertainty (TPU) as a Motive for Trade Agreements

In [Section 2](#), we described the depth of policy cooperation in PTAs goes beyond currently applied tariffs and NTBs. It also includes provisions about future policies such as tariff bindings and contingent protection. In this section, we review recent theory and

^{zz} The basic approach is to attribute the purely bilateral trade differences predicted by the model relative to the data to a trade wedge, converted to an ad valorem equivalent using a trade elasticity, 4 (so similar in spirit to what is discussed in the NTB section). This measure is then regressed on PTA indicators, a bilateral pair and country-by-year effects, as we did in [Table 3](#) for tariffs.

evidence that suggests that PTAs reduce expected protection and uncertainty about future policy and that, by securing future market access, PTAs increase current trade-related investments and trade volume. I will also argue how certain types of TPU models can help explain the heterogeneous trade effects of PTAs.

To understand the impacts of deeper PTAs we must model and estimate the effects of specific policies at a disaggregated level. This subsection describes how this has been done in the context TPU.

4.3.1 Sources of TPU

A reasonable starting point for examining the motives for deeper PTAs is to ask what their stated goals are. One of them is for PTAs to “ensure a predictable environment for business planning and investment”; as stated in several agreements undertaken by the United States, European Union and several developing countries.^{aaa} This motive is present in other trade agreements, for example the WTO states that “Just as important as free trade—perhaps more important—are other principles of the WTO system. For example: nondiscrimination, and making sure the conditions for trade are stable, predictable and transparent.” Despite these stated objectives, until recently the literature on agreements mostly treated trade policy as deterministic. It is thus important to understand why and to point out some potential sources of TPU.

One possible reason why TPU has been underresearched is the perception that trade policy is not very volatile; after all statutory tariff rates are legislated at most on a yearly basis. However, this perceived low volatility in statutory tariffs is a misleading guide for the degree of TPU for two reasons. First, even if statutory trade reforms are infrequent when they occur the changes can be quite large and persistent, as documented by [Bown and Crowley \(2016\)](#). Second, applied trade policy is more volatile than statutory tariff rates due to NTBs that are not strictly regulated by the WTO. While some of these are meant to be “temporary” they can remain in place for months or years (cf. [Bown, 2011](#)).

The ability to use unregulated trade policies can interact with macroeconomic or political shocks to generate considerable uncertainty. For example, there was widespread fear that the 2008 economic crisis would result in a substantial increase in protectionism. This included the possibility of antidumping measures, increases in developing country tariffs from their applied level to the maximum allowed under the WTO, and the use of government procurement measures to favor national firms. These fears were reasonable since there is evidence that protection responds to a variety of economic shocks such as aggregate downturns (cf. [Bown and Crowley, 2013](#)). Even though the worst fears of a trade war were not realized, its possibility created uncertainty, as evidenced by governments’ repeated assurances that they would not resort to 1930s type protectionism.

^{aaa} For examples see the texts in Global PTA Database at <wits.worldbank.org/GPTAD>.

The possibility of a rare event such as trade war should only be a concern if it leads to very high protection. That was clearly the case in the 1930s and a clear example remains in the form of US column 2 tariffs that are applied to certain non-WTO members and are on average 35%. Moreover, there is evidence that countries have substantial import market power and explore it when they are not bound by agreements such as the WTO (Broda et al., 2008); this incentive is less pronounced when they accede (Bagwell and Staiger, 2011) but does not disappear (Ludema and Mayda, 2013). Finally, Ossa (2014) estimates optimal tariffs for individual countries in 2007 and the median is about 60%.^{bbb}

Another possible source of TPU is domestic political shocks, eg, due to changes in government or lobbying pressures. Amador and Bagwell (2013) show that if governments have private information about those shocks, there are contracting imperfections and terms-of-trade externalities then it would be optimal for an agreement such as the WTO to impose tariff bindings. Such an agreement lowers TPU and increases trade. More broadly, agreements such as the WTO can also increase transparency and thus the degree of trade policy cooperation in a repeated game.^{ccc}

While WTO accession may lower uncertainty about future protection, it does not eliminate it. There are ongoing negotiations and when they are finalized there is uncertainty in terms of implementation and the possibility of substitution toward unregulated policies, particularly in times of economic crisis. Moreover, there are issues on which its members have not agreed and have the potential to trigger disputes and high protection. These include (i) quality and safety concerns that raise the possibility of product bans; (ii) the US threat of import duties to counter Chinese currency manipulation; and (iii) the possibility of using “environmental” duties at the border to offset differences in carbon emissions in production.

Deeper PTAs can reduce TPU arising from some of these sources. First, to the extent that certain barriers are eliminated and bound at zero there is less risk of future renegotiation. Second, countries with more similar preferences may be able to agree on standards and various noneconomic policies and include mechanisms for cooperating in them and reduce the risk of them triggering future protection. More generally, it seems

^{bbb} Authors calculations based on the median estimate across all industries and countries.

^{ccc} Beshkar and Bond (2015) show how uncertainty coupled with contracting imperfections can explain the use of tariff bindings. Beshkar et al. (2015) model the relationship between tariff bindings and applied tariffs in the negotiation of agreements and study their empirical relationship with importer market power. In certain cases, it may be optimal for the agreement to allow for policy variability as a way to sustain cooperation in periods when the incentive to defect in a repeated game is high, eg, if there are terms-of-trade motives for protection and import volumes increase (cf. Bagwell and Staiger, 1990). These studies typically focus on the optimal policy response to shocks and uncertainty so we briefly return to them in the context of endogenous policy formation in Section 5.

plausible that certain PTAs can reduce TPU by (i) securing low (often zero) tariff rates that are fixed over time and less subject to being eroded via policy substitution (either temporary, eg, antidumping, or permanent, eg, product standards) and (ii) integrating the economies to make a trade war extremely costly. We now discuss whether there is evidence for these hypotheses.

4.3.2 Direct Evidence of TPU Reductions via Agreements

There are different dimensions of TPU. Some of these are directly observable, eg, tariff volatility over time, how frequently contingent policy is used, how frequently a policy regime is reviewed and/or canceled. In this section, I review the evidence of the impact of PTAs on each of these dimensions. Other dimensions of TPU are harder to detect and measure and addressed in the next subsection.

Two pieces of evidence suggest that agreements lower volatility in trade policy. First, tariffs in developed countries may not be very volatile simply because they have long been members of trade agreements or because their tariffs are never volatile. To address this issue, [Limão and Maggi \(2015\)](#) examine the average US tariff over pre- and post-agreement years, namely from 1860 to 1960. The standard deviation of that policy (over time) before 1934 is at least twice as high as during 1934–61—a period marked by the Reciprocal Trade Agreement Act (1934) and the signing of GATT (1948). They note that the higher volatility of US trade policy before 1934 reflects the Smoot–Hawley tariff hike of 1930 but also several prior major changes in the tariff code. Second, [Cadot et al. \(2010\)](#) provide econometric evidence of the impact of PTAs on policy volatility. They use a panel of changes in price distortions introduced by agricultural trade policies and find that the absolute value of that change falls when countries enter PTAs.

As countries lower tariffs in PTAs they may use contingent protection to deal with particular shocks. By this measure PTAs could increase TPU toward members. [Prusa and Teh \(2010\)](#) examine this question and find the opposite. They estimate that antidumping provisions in PTAs lowered AD cases between members by as much as 55%. But they find PTAs increased AD toward nonmembers, which suggests a TPU externality that we discuss in [Section 6](#).

An alternative to examining the impact of accessions to an agreement is to examine if adoption of specific parts of it affect TPU. [Grosso and Piermartini \(2014\)](#) for example examine the impact of WTO tariff bindings on the probability of MFN tariff increases. Using HS-6 data from 1996 to 2001 for all WTO members they find that bindings reduce the probability of increases in the MFN applied rate.

It is important to note that not all PTAs reduce TPU and in fact some may increase it. For example, unilateral preferences such as the GSP provide recipients with tariffs below MFN but these are subject to renewal and cancellation risk by the “donors” (the United States, and other developed countries). Temporary and permanent cancellations do occur

and have negative impact on recipient exports.^{ddd} This has been recognized and one of the objectives in the recent reform of the GSP by the European Union was to put it in force for a longer period, 10 years instead of 3, to “enhance stability and predictability” and “improv(e) certainty for business operators.” This type of change provides an example of an increase in PTA depth that reduces TPU. [Ornelas \(2016\)](#) provides additional discussion of GSP uncertainty for its recipients.

Similarly to GSP, other unilateral preferential programs are subject to renewal and cancellation. This has lead countries to seek deeper, reciprocal agreements to extend, and secure preexisting preferences. Two examples are Peru and Colombia, which sought and obtained reciprocal PTAs with the United States and argued this security would be important for export investments (cf. [USITC, 2008](#)). There are other examples, one of which we will examine later.

A final point to note is that even if we do not observe any volatility in a given policy over a certain period this does not imply there is no TPU. If exporting firms believe that a sufficiently large shock would change future trade policy they will take this information into account. Whether and how exactly they do so depends on the mechanism linking TPU and their investment decisions, which we now describe.

4.4 A TPU-Investment Mechanism

4.4.1 Mechanism

One obvious channel through which future trade policy can affect current trade values is via firm investments in the tradable sector. However, whether foreign TPU increases or decreases a country’s exports to that market is not obvious. We briefly describe why and then focus on an option value mechanism that generates a negative impact of TPU on export investments and trade. The mechanism is tractable and seems to capture the concerns voiced by businesses in the context of trade policy.

[Limão and Maggi \(2015\)](#) use a standard general equilibrium model and ask under what conditions governments would choose to form an agreement to reduce TPU. They show that TPU lowers investment and trade only if there is sufficient income risk aversion. This is a necessary condition to overcome a basic force present in most standard models when agents make ex ante decisions based on expected values. Suppose for example that a firm must make a once and for all decision on whether to invest in an export-related activity based on its expected value. A mean preserving spread of the product price increases that expected value if the firm can adjust any of its inputs ex post to take advantage of a price change. This convexity of profits with respect to product prices implies that increases in foreign tariff risk can actually increase exports unless there is sufficient income risk aversion.

^{ddd} See trade.ec.europa.eu/doclib/docs/2013/december/tradoc_152015.pdf, which also notes that in 2014 the European Union canceled 86 of the 176 GSP beneficiaries.

Handley and Limão (2015) explore an alternative mechanism that generates a negative relationship between TPU and exports. There is evidence that exporting requires sunk investments that are at least partially irreversible. If policy is sufficiently persistent then firms can wait to observe the policy conditions and invest only if they are sufficiently favorable. Thus TPU generates an option value for export-related investments and reduces the mass of exporters and export value. Handley and Limão (2013) extend the mechanism to productivity enhancing sunk cost investments and show that TPU can thus also reduce exports of incumbent firms.

To understand some of the evidence below it is useful to outline the basic elements of their model. After any export-related investments, firms choose production to maximize operating profits subject to a CES demand as in a Melitz–Chaney model with heterogeneous productivity. Thus the export investments are the only decisions made under uncertainty. To do so firms observe current policy and the policy regime (the probability it will change and the distribution of values if it does) and invest if the present discounted value of doing so net of the sunk cost exceeds the optimal value of waiting until the foreign barrier is lower or less uncertain. The optimal stopping problem generates the following marginal cost cutoff below which all firms from x enter m at t ,

$$c_{xmt}^u = U(r_{xmt}, \tau_{xmt}, \tau_{xmt}^{hi}) \times c_{xmt}^d$$

where c_{xmt}^d is the cutoff in the absence of TPU and $U < 1$ is the uncertainty term, which implies less entry under uncertainty.

The uncertainty term reflects the exporting firms' belief that with probability r_{xmt} the current policy in the export market, τ_{xmt} , will increase to some value τ_{xmt}^{hi} . An increase in uncertainty measured by higher probability that the policy will increase or decrease relative to the current value implies a higher r_{xmt} and a lower expected value of current entry even if the expected value of the tariff is unchanged. This is an example of the bad news principle and reflects the fact that if conditions improve the firm can enter and take advantage of it, but if it is already in and conditions deteriorate it suffers a profit loss.

The model generates a functional form for U that reflects the proportional profits lost conditional on a bad shock, which is a function of the current tariff and a counterfactual value τ_{xmt}^{hi} . Since the deterministic cutoff takes a form similar to Chaney (2008) we can combine these to write a TPU augmented gravity equation.

4.4.2 TPU Augmented Industry Gravity

This mechanism allows us to study the impact of TPU and PTAs that may affect it on different measures of exports and related investments. Here we briefly discuss how two of them can be analyzed using an industry gravity approach. We first focus on export values to easily build on and compare with Section 3. We will then point out the relationship with number of firms and resulting entry investments.

4.4.2.1 Export Values

Since the TPU mechanism provides an expression of the marginal firm in terms of TPU we can use a particular productivity distribution, say Pareto, to aggregate firm export sales for a given industry V and derive a theory consistent TPU augmented gravity equation. Using the structural gravity notation in Section 3, we can write it as follows:

$$\ln T_{xmtV} = \ln \phi_{xmtV} + \alpha_{xt} + \alpha_{xV} + \alpha_{mt} + \alpha_{mV} + u_{xmtV} \quad (11)$$

$$\ln \phi_{xmtV} \equiv \varepsilon_U \ln U(r_{xmt}, \tau_{xmtV}, \tau_{xmtV}^{hi}) - \varepsilon_\tau \ln \tau_{xmtV} - \varepsilon_d \ln d_{xmtV} + \alpha_{xmV}$$

There are two basic differences between (11) and the aggregate specification in (7). First, (11) applies to each industry so there is an additional dimension of variation. The model imposes some structure and yields multilateral terms, $\alpha_t + \alpha_V$, that are log separable into aggregate and industry-specific components.^{eee}

The more fundamental difference between (11) and (7) is the bilateral market access function. Instead of estimating the partial PTA effect as an average treatment, (11) models changes in current and future policy. In principle we can use any observable policy with bilateral effects, but the existing research has focused on different types of tariffs and thus so will we. In the absence of policy uncertainty (or conditional on it) the partial elasticity of trade with respect to applied tariffs is $\varepsilon_\tau > 0$. If exporters believe there is a risk, measured by the probability r_{xmt} , that protection will increase from its current level to τ_{xmtV}^{hi} then there will be lower investment in export-related activities and $\ln U < 0$ so exports are lower, therefore $\varepsilon_U > 0$. Note also that if $\ln U = 0$ then (11) reduces to a standard industry gravity equation, hence the TPU augmented gravity.

The bilateral access function also controls for trade costs not included in the policy terms. Namely, it includes observed changes in trade costs (insurance and freight in d_{xmtV}) as well as unobserved factors, α_{xmV} . The latter avoids the type of bilateral selection bias discussed in Section 3 and implies that the identification is obtained from changes in applied policy and uncertainty surrounding it. How we measure uncertainty depends on the specific setting as described later.

In Section 3, we provided evidence that at reasonable values of current estimates of trade elasticities $\tilde{\varepsilon}$ the observed tariff changes in PTAs were typically too low, or alternatively, the required elasticity given those tariffs, too high. From (11) it should be clear that one way that TPU can help explain this is by taking into account changes in future protection, ie, adjusting for the fact that exports before the PTA were depressed. A second and less obvious implication is that if TPU is present but ignored then estimates of the tariff elasticity are biased downwards. This attenuation effect is simply due to the fact that under uncertainty some of the current tariff change is expected to be reversed. Thus if ex ante we used such tariff elasticity estimates to predict the impact of a given tariff

^{eee} These capture exporter aggregate cost shocks and productivity parameters, importer expenditure shocks and changes in multilateral resistance terms.

change in a PTA we would under predict the true outcome if the PTA also lowered uncertainty. To the extent that countries with lower tariffs may also have lower uncertainty, our earlier finding that the tariff elasticity is higher at lower levels may reflect an uncertainty effect. But we can test the mechanism directly.

4.4.2.2 Firms/Varieties

If we are interested in the impact of the PTA on export investments or the number of exporting firms (or varieties) we can also employ the approach just outlined by using the number of firms (or varieties) as a dependent variable. Moreover, under a Pareto distribution the interpretation of the parameters is unchanged except that the structural interpretation of ε_U is different, and we expect it to be smaller in the export equation because entering firms are smaller on average.

4.5 Ex Post Trade and Firm/Variety Entry Investment Estimates of Deeper PTAs

We now describe the application of this framework in different settings. To illustrate the main points we focus on a specific application and then briefly list the others.

4.5.1 Expanding and Securing Existing Preferences

Some PTAs that involve nonreciprocal and/or temporary preferences can leave exporters with considerable uncertainty about future market access. We discuss examples in the introduction of this section. One other example was Portugal's access to the EC and Spanish markets prior to 1986. As part of EFTA Portugal enjoyed duty free access to the EC in industrial products since 1977 and faced Spanish tariffs that were about half of that country's MFN tariff since the early 1980s. After accession these tariffs went to zero and importantly were no longer expected to change.

Handley and Limão (2015) show that even focusing on the years immediately after accession there was extremely strong entry of Portuguese exporters to serve those markets. Using Portuguese exports and the aggregate gravity approach in (7) they estimate a partial effect, $\tilde{\phi}$ of 23 log points for the EC and 115 for Spain. The EC tariff reductions were minimal so clearly something else must explain the effect. Spain's tariff reductions were on average 6.6 log points, which require a tariff elasticity of 17.4 to account for the effect—a clear example of the elasticity puzzle.

These authors then analyze to what extent the puzzle can be explained by the reduction of TPU, ie, the fact that accession may have permanently secured preexisting preferences and/or reduced applied protection. To do so they apply the approach in (11) and measure the proportion of profits lost for any given firm in industry V conditional on losing the preferences as $1 - (\tau_{xmtV} / \tau_{mtV}^{MFN})^\sigma$ where x is Portugal, τ_{xmtV} is the applied tariff it faces in market $m = \text{Spain}$ or any of EC-10 countries and τ_{mtV}^{MFN} is the counterfactual that

it would face if it lost the preferences, which they take to be those countries MFN tariffs on GATT members. For a given σ we can construct this measure and estimate (11) where the parameter on this variable is time varying only to the extent that firms change their belief about the probability of losing the preferences. So if $r_{xmt} > 0$ prior to the agreement and it falls or is zero after this indicates that the PTA reduced TPU.

Using firm-level exports they estimate the effect of accession on industry net entry by Portuguese firms and their total exports to the EC and Spain. They find evidence of a positive probability of reversal before 1986 but not after. The combined effect of changes in TPU and applied tariffs accounted for 61% of the observed firm entry growth and 87% of export growth in the period examined.

4.5.2 Average Treatment vs Policy Effects

To understand the relation of the TPU gravity with the aggregate partial effect we consider the findings for exports to Spain. Recall that the aggregate partial effect for Spain was $\tilde{\phi} = 1.15$. In any given industry the combined policy effect if TPU is removed is given by

$$\tilde{\phi}_V^{tpu} = -\varepsilon_U \ln U_V - \varepsilon_\tau \hat{\tau}_V.$$

So the average aggregate effect is simply a weighted average of these, which is equal to 85 log points for Spain. If tariffs had not changed then the only contribution would come from the uncertainty term, ie, from securing preexisting preferences, which is estimated to be 20 log points.^{fff} A less obvious impact of eliminating TPU is that it increases the elasticity of any given tariff change by ensuring that it will not be reversed. They estimate that if tariffs had been reduced but uncertainty had not then exports would have grown by 45 log points. Thus the remaining 20 log points represent the role of TPU reduction in locking-in those tariff reductions. In sum, if the accession had only lowered Spanish tariffs then we would be able to account for less than 40% of the average treatment effect (45/115) and accounting for TPU brings this number up to about 75%.

4.5.3 PTAs as Insurance Against Trade Wars

After the 2008 financial crisis, trade fell much faster than income worldwide, a puzzle for standard trade models, and typical estimates of the income elasticity of trade. The current explanations for this episode can account for some of the trade decline but ignore the subsequent fast recovery. More importantly, they are silent about the impacts of the potential trade war that was feared by policy makers.

To place this event in perspective, note that in 2009, world trade fell by 12%, the largest decline since the great depression (10%), but income fell only 2.7% vs 20% for

^{fff} Evaluating this change at the long-run mean tariff yields the pure mean risk reduction effect, which is estimated to be about 73% of the uncertainty reduction at initial tariffs.

industrial output in depression (Eichengreen and O'Rourke, 2009). Moreover, the WTO (2011) shows that applied trade barriers affected only 1% of trade and Kee et al. (2013) estimate these accounted for less than 2% of the collapse. In contrast, trade barriers in the great depression increased by as much as 35% for the United States, Germany, and France and accounted for large fraction of decline according to Madsen (2001).

One important difference relative to the Depression is the current network of trade agreements. These include the GATT/WTO, which was created in response to the 1930s trade war to prevent a reoccurrence and this role was noted during the crisis.^{ggg} While this institution did not fully eliminate the *possibility* of a trade war its monitoring mechanism may have helped prevent it from realizing. It is also possible that the extensive network of PTAs may have reduced the probability of a trade war between partners. The question is whether and how we may test these hypotheses and quantify their implications.

Carballo et al. (2015) examine the role of PTAs as insurance against trade wars. They extend Handley and Limão (2015) to encompass an exit margin and allow for demand uncertainty both from policy and income shocks, and they further allow for policy uncertainty to increase as a result of higher income uncertainty as suggested by the great trade collapse (GTC). The model has several predictions. The key ones are that if a PTA reduces TPU then:

- (i) The increase in income uncertainty during the GTC would have lowered exports to non-PTA markets by relatively more (since in PTA markets there would not been much of an increase in TPU) and this effect is stronger for the extensive margin, ie, relatively more net exit of firms/varieties from non-PTA markets.
- (ii) In non-PTA markets the effects in (i) are stronger in industries where market power is higher so exporters fear higher losses in case of a trade war. But in PTA markets there should be no differential effect across industries if there is no probability of a trade war.

To test these predictions they use US firm-level data. They first establish that the extensive margin accounts for 43% of the collapse in US exports to non-PTA but only 28% for PTA (in Q3–08 to Q3–09). They then construct measures of income uncertainty in export markets and estimate if these had differential impacts on US exporters' entry and exit over time, PTA membership, and market power. They find this measure of uncertainty leads to a reduction in varieties that was 10 percentage points higher for non-PTA markets than PTAs between Q4–08 and Q3–10. This differential disappeared

^{ggg} In 2009 its Director General Pascal Lamy stated that "Today as the economic crisis bites into our economies, and as protectionist pressures knock on our doors, we must recall the importance of the insurance policy against protectionism that the WTO offers through 60 years of global rule-making, and its dispute settlement system" www.wto.org/english/news_e/sppl_e/sppl112_e.htm.

after Q4–10 when it was apparent that no trade war was imminent. Furthermore, the reduction in varieties exported to non-PTA was larger for relatively higher market power industries.

The relatively lower net exit from PTA markets caused by economic uncertainty translated into relatively higher exports during the crisis.^{hhh} Thus the current evidence suggests PTAs can play an important role in insuring against potential trade wars. Further analysis would be important in estimating and quantifying this channel in other settings.

4.6 Other Evidence and Future Work

There is also evidence that the WTO increases trade via reductions in TPU. [Handley \(2014\)](#) uses panel data for Australia and finds that higher uncertainty, as measured by the gap between applied MFN and bound tariffs lowers the probability of importing an HS-8 product. [Deason \(2014\)](#) applies his methodology to a broad cross section of countries at the HS-6 level and finds similar results for the typical country in the sample.

TPU also provides a possible explanation why PTAs can have heterogeneous effects. If firms do not believe the current policy changes are credible then their response will be attenuated. Therefore, the depth of PTAs, as measured by the credibility of the provisions and the presence of enforcement mechanisms is critical in generating investment and trade effects.

The potential heterogeneity in PTAs indicates there is a high value for future research of specific agreements, which should take into account actual policies and their potential worst case counterfactuals. Doing so with aggregate data is not feasible because of aggregation bias and the small sample issues described in [Section 3](#). However, we can explore detailed product and firm-level data and a particular framework to aggregate the results if so desired. This section also highlights the importance of using dynamic models to better understand and estimate the impacts of PTAs both their formation and how they transmit shocks.ⁱⁱⁱ

One important policy implication of the research on TPU is that PTAs can have large trade effects even if tariffs and NTBs are relatively low. The [WTO \(2011\)](#) pointed that only 16% of world trade took place under positive preferential margins and interpreted this to mean that the motive for current PTAs must no longer be tied to reducing tariff barriers. This ignores the fact that PTAs can reduce uncertainty about future protection and act as an insurance against trade wars. Thus even if current tariffs are low PTAs can still have important discriminatory effects against nonmembers, which we discuss in [Section 6](#).

^{hhh} This can help to explain why the share of US exports to PTAs, which had declined almost 5 percentage points between 2005 and the start of the crisis, stabilized and increased moderately since 2009.

ⁱⁱⁱ [Ruhl \(2004\)](#) argues that PTAs can generate export entry by permanently lowering trade frictions, which strengthens response to future macrovariable shocks.

5. PTA FORMATION AND POLICIES: MOTIVES AND DETERMINANTS

Thus far we summarized key features of PTAs and their effects on trade and related policies. In the process, we discussed some of the stated objectives of PTAs and how accounting for their endogenous formation affects the estimation of their impacts. This section focuses on the determinants of the formation of PTAs and of preferential tariffs.

The starting point is the motives for traditional static PTAs and the evidence that focuses directly on the mechanisms behind them: trade diversion and terms-of-trade effects. I then describe some nontraditional motives for PTAs and evidence for some mechanisms underlying them, which have received less empirical attention. In [Section 5.2](#), I discuss the empirical determinants of (i) PTAs between pairs of countries and (ii) endogenous preferential tariff levels.

5.1 Motives and Mechanisms

5.1.1 *Traditional*

Most empirical work on the economic determinants of PTAs focuses on trade-related motives. We start by describing the traditional motives for forming a PTA and the evidence for the mechanisms behind them such as trade diversion and price effects.

5.1.1.1 Trade Creation and Diversion

The main question that the traditional analysis of PTAs asks is if exogenously lowering tariffs between a pair of countries increases the social welfare of members and nonmembers?ⁱⁱⁱ The answer depends on the economic structure as well as the pattern of initial and final tariffs. The possibility that the elimination of *distortionary* tariffs between two members may reduce their own welfare would appear to be a special case. But due to [Lipsey and Lancaster \(1956–57\)](#) we now understand it is a basic example of the principle of the second best: if the initial tariffs are not at their unilateral optimum then an exogenous change in them will generally have an ambiguous welfare effect.

One of the contributions of the traditional analysis of PTAs is to show under what conditions exogenous PTAs generate net welfare losses for members and nonmembers. The resulting insights have provided a guide for the initial empirical analysis of the economic determinants of PTA formation. The basic trade off that arises in many of these models is that a PTA lowers the cost and distortion from the PTA member and in doing so leads to trade creation, but it can also generate trade diversion: a substitution away from the nonmember. [Viner \(1950\)](#) shows that if the nonmember is the lower cost supplier then the trade diversion cost can offset the gain from trade creation.

In [Section 3](#), we provided evidence of positive partial PTA effects between members. What do those theoretically consistent gravity estimates by themselves tell us about

ⁱⁱⁱ See [Krishna \(2008\)](#) for a review of the related literature.

diversion vs creation? Not much. Recall from Eq. (5) that the partial PTA estimate identifies the increase in member trade relative to nonmembers. So the 60 log point increase in Table 2 may be all due to creation and no diversion or the opposite or any combination of the two. If in addition we could estimate the aggregate impact of the PTA for a given importer we could determine diversion but that aggregate impact is subsumed in the importer fixed effects.^{kkk}

We describe two approaches to explore bilateral trade flows to determine the extent of trade diversion. The first takes advantage of a structural model to disentangle the two effects. Egger et al. (2011) estimate the average partial PTA effect, translate it into an ad valorem equivalent and implement the counterfactual of no PTAs in the context of an Armington model. They compute the general equilibrium trade of PTA members with nonmembers, which falls by 9% on average with declines for about two-thirds of the pairs. This is a modest decline relative to the average trade creation between members of 102%. This suggests that on average PTAs are creating more trade than they divert. However, given the heterogeneity of effects across countries it would be interesting to calculate this net trade effect by country as well.

The second approach is to explore additional sources of data variation. Clausing (2001) finds little effect of CUSFTA on US-Canada trade using an aggregate gravity approach but finds significant increases in US imports from Canada when using detailed commodity data and tariff changes and using year dummies to control for aggregate effects. There is no evidence of diversion: the US import share from the rest of the world does not fall by more in products with higher preferential reductions.^{lll} In contrast to Clausing (2001), Romalis (2007) does find substantial trade diversion due to the CUSFTA (and NAFTA). The main methodological distinction is that the latter study uses differences-in-differences: it examines US imports from nonmembers *relative* to those of the European Union and finds larger relative import reductions in HS-6 goods that obtained larger preference margins. Trefler (2004) also employs disaggregated tariff changes but focuses on Canadian imports and finds trade diversion.

5.1.1.2 Price Effects

We now turn to an alternative outcome—price changes—which can be informative about the net changes in trade and welfare due to a PTA. We first describe the link between trade diversion, prices, and welfare for a small country and then consider more generally the role and evidence of terms-of-trade effects in the context of PTAs.

^{kkk} Recall that these are used to control for multilateral resistance—something that is omitted in earlier studies that attempt to estimate trade diversion, thus we do not discuss those studies.

^{lll} Magee (2015) follows a similar approach and finds that Turkish imports did not exhibit much trade diversion as a result of its CU with the European Union.

How much diversion is required to generate welfare losses for members? In a setting where trade diversion from a nonmember is necessary for a member to lose from a PTA then if they do not trade there can be no diversion or net cost from the PTA. This extreme case of no trade with nonmembers is the basis for the natural trading partner hypothesis (cf. [Krugman, 1991](#)) that claims that the welfare gains from PTAs are expected to be higher if the partners are “natural,” ie, trade mostly with each other before the PTA.

[Krishna \(2003\)](#) examines if there is direct evidence for the natural trading partner hypothesis by exploring price data. He estimates the welfare effects for the United States of a unilateral preferential tariff reduction toward alternative countries. Assuming the United States is small and there is an Armington structure the welfare effect is equal to a weighted average of the bilateral trade created with the PTA member and the amount diverted from nonmembers, where the weight is the initial tariff faced by each.^{mmmm} In this simple setting, if initial tariffs are identical, we could add up the estimated trade effects to determine welfare for the member. Instead [Krishna \(2003\)](#) estimates the relevant own and cross-price import demand elasticities. The preferential reduction reduces the US *consumer* price for the PTA import and the own price elasticity estimates imply this creates trade but the cross-price elasticities also show there is substitution away from the nonmembers (so diversion). The cross-price effects are sufficiently small that US welfare would increase if it reduced its tariff preferentially with respect to any one of the 24 countries considered. These *ex ante* gains from *potential* PTAs are not correlated with distance and thus he concludes there is no evidence for the natural trading partner hypothesis.

One of the most commonly examined sources of economic policy externalities in the trade setting is the terms-of-trade externality. Such an externality is present if the price that exporters from x receive in m , $p_{xm}(\tau_{mx},)$, depends on the latter’s tariff. When the initial tariff is not set cooperatively there will generally be some τ'_{mx} that improves the objective in country x so x has a bilateral policy externality motive for a PTA with m . When there are only two countries this is not only one of, but often the only externality that standard agreements over tariffs address (cf. [Bagwell and Staiger, 2016](#)).ⁿⁿⁿ

The bilateral terms-of-trade motive for trade agreements is well understood in the context of two countries. The marginal gain for an exporter x from facing lower tariffs in m reflects increases in its export price from increased market access to m . The potential cost of the agreement for x (if its initial unilateral tariff is optimally set) reflects its terms-of-trade motive for a tariff, which when reduced deteriorates its bilateral terms-of-trade.

^{mmmm} [Trefler \(2004\)](#) applies this criterion to argue Canadian welfare increased due to its tariff reductions on US goods.

ⁿⁿⁿ That is the case not only in settings where the government objective reflects social welfare in a perfectly competitive setting but also when it reflects political economy considerations and in certain noncompetitive environments.

If we only have two countries then any bilateral terms-of-trade gain by x imposes larger losses on m and so some bilateral tariff reduction between them would be optimal. Thus, all else equal, the bilateral terms-of-trade motive predicts PTAs are more likely between countries with relatively higher bilateral import market power. To my knowledge, this has not been directly tested as a PTA determinant but below we provide a suggestion for doing so.

We can also relate the terms-of-trade externality to the trade diversion effects of PTAs. Consider a preferential tariff reduction between x and m while keeping other policies fixed. Assume some substitution in consumption between member and nonmember goods. With three countries we need to consider how the bilateral tariff affects the multilateral terms-of-trade. The bilateral tariff reduction generates substitution and thus increases the *share* of imports from the PTA partner (as estimated by the partial average PTA effect); moreover, if the PTA member is less efficient than the rest of the world then this reallocation toward the relatively more expensive good implies a deterioration of home's terms-of-trade.^{ooo}

We can thus provide a terms-of-trade interpretation of the trade diversion effect, which has two advantages. First, to determine the welfare effects of PTAs based on a trade diversion criterion we need to estimate if it occurred, which is not straightforward as discussed earlier, so having another measure, changes in terms-of-trade, to examine this potential cost of PTAs is useful. Second, even if diversion is present it is not sufficient to infer welfare effects whereas in certain models we can directly relate price changes to welfare. For nonmembers, a decline in their terms-of-trade will often be a sufficient statistic for whether they are harmed (cf. [Winters, 1997](#)). For members, if their terms-of-trade relative to the rest of the world improve then this is an additional benefit to internalizing their bilateral terms-of-trade externality. Some of these points can and have been formalized and we return to them in [Section 6.1](#) when discussing the existence of necessarily welfare-enhancing PTAs.

There is both direct and indirect evidence for the impact of tariffs on prices. The direct evidence examines the impact of tariffs on either (i) the same good's price, $\frac{\partial p_{xm}}{\partial \tau_{mx}}$, or (ii) on the price received by a nonmember, $\frac{\partial p_{um}}{\partial \tau_{mx}}$. The first effect is the standard tariff pass-through effect and evidence for it was pioneered by [Kreinin \(1961\)](#), which finds that US reductions in its multilateral tariffs lead to increases in the prices received by exporters to the United States in two-thirds of products, so $\frac{\partial p_{xm}}{\partial \tau_{mx}} < 0$.

[Chang and Winters \(2002\)](#) provide evidence for PTA price effects on nonmembers. They examine the impact of Mercosur's preferential tariff reductions on the relative

^{ooo} So even if home is small so it does not affect the price received by other countries, the tariff reduction can still increase its pretariff average import price if it shifts demand to the costlier supplier.

prices received by nonmember countries exporting to Brazil. They derive a pricing equation from a Bertrand game between nonmember and member firms similar to the one [Feenstra \(1989\)](#) uses to study US tariff pass-through. In this setting, a reduction in Brazil's tariffs on Argentina shifts demand away from nonmembers' firms and leads them to adjust the price down if their perceived demand becomes more elastic. This is what they find using unit values for detailed product data. Similarly, [Winters and Chang \(2000\)](#) find that when Spain enters the European Union the relative price received by non-European Union exporters to Spain fell. So, both studies provide direct evidence that PTAs generate a negative terms-of-trade externality for nonmembers.

The indirect evidence of PTAs on terms-of-trade infers price and/or welfare changes after using bilateral trade flows to estimate price elasticities and then evaluating the effect of tariff changes or PTAs in the context of a model. We discussed these briefly at the end of [Section 3](#), here we expand on the findings of [Anderson and Yotov \(2016\)](#).^{PPP} They provide a variety of counterfactual effects from removing all PTAs that took place in 1990–2002. They calculate the AVE of PTAs as defined in (8) from their gravity estimates and use this as the shock in a multicountry Armington model with eight manufacturing sectors. They compute terms-of-trade effects as the ratio of a country's aggregate seller price index relative to the country's import price index. The latter reflects trade costs so it is possible for terms-of-trade, as they define them, to increase for all countries. Relative to a counterfactual with no PTAs they find that PTAs improved the TOT for most of the countries with the few losers experiencing small losses. The import price index fell for all countries, particularly small eastern European countries entering the European Union. The export price effects of PTAs on the other hand were more variable—falling in about half the countries—with larger losses for non-PTA countries. This highlights the multilateral externality of PTAs in lowering nonmember export prices.⁹⁹⁹

Anderson and Yotov also perform alternative counterfactuals eliminating a couple of specific PTAs. It would be interesting to compute counterfactual effects of PTAs for all bilateral pairs (one at a time or in different combinations) whether or not they had a PTA. This could then be used to evaluate if the PTAs that did form were the ones that yielded higher bilateral welfare.

^{PPP} [Romalis \(2007\)](#) also evaluates aggregate effects of NAFTA. He estimates an average import and export supply elasticity for the United States and uses them to simulate a welfare change due to the observed tariff changes based on an aggregate CES price change for given production net of tariff revenue. He finds that the aggregate price gains for members are almost fully offset by lost tariff revenues, an indication of the diversion effect this same paper finds. Because he focuses only on applied tariff changes in NAFTA the estimates only capture a fraction of the trade effect.

⁹⁹⁹ The overall TOT gains are an overestimate of the income effect in this setting because they ignore tariff revenue; the authors argue the latter diminishes but does not overturn the conclusion.

5.1.2 Nontraditional

The traditional view of PTAs as an exogenous preferential tariff reduction is a modeling abstraction, as is clear from the variable depth and breadth of economic policies that modern PTAs address. Focusing solely on the social welfare costs relevant in the traditional view and ignoring nontraditional motives and economic structures for PTAs has two potential costs. First, we may miss important outcomes of PTAs, not just on trade (as shown in [Sections 3 and 4](#)) but on FDI, and other outcomes. Second, the traditional view restricts our attention to a single binary policy outcome: an exogenous tariff reduction between a pair of countries. If we do so then we fail to explore a rich set of alternative PTA policy outcomes, including endogenous tariff variation, and also ignore or misinterpret the empirical effects of determinants of the probability of a PTA forming.

There is an extensive set of nontraditional settings to consider that reflect both economic and noneconomic objectives. Our goal here is to highlight a few and describe what if any evidence there is suggesting their potential importance. Let us first more precisely define the “traditional view of PTAs” and then characterize the alternatives to it as deviations either in terms of the objective or economic setting. PTAs are inter-governmental agreements and so their formation depends on what we specify as the government’s objective, G_x . The traditional view effectively places three restrictions on G_x : (R1) it is a measure of social welfare; (R2) it depends only on applied tariffs; and (R3) policy levels are exogenously given in an agreement. In contrast, the nontraditional motives discussed later allow G_x to reflect redistributive or other political economy considerations over many policies that it sets endogenously. Moreover, the traditional view focuses on restricted settings that: (R4) allow only for trade in goods; (R5) are static; (R6) exclude bargaining or enforcement considerations; and (R7) ignore nonpecuniary externalities. Different nontraditional settings relax these restrictions as follows.

5.1.2.1 Endogenous Trade Policy

There is substantial evidence that tariffs and other trade policies are endogenous to political economy and economic factors and many of these are incorporated in the PTA research since the 1990s.^{xxx} In the context of PTAs the WTO members do face some constraints on preferential tariff levels, but there are numerous exceptions that countries can and do build into PTAs (cf. [WTO, 2011](#); [Bown and Crowley, 2016](#)). Even if the WTO enforced the Article XXIV constraint that PTAs set most preferential tariffs at zero, the initial tariff could still reflect terms-of-trade and political economy motivations.

There are two basic implications of relaxing the exogenous tariff assumption, R3, even if G_x represents social welfare. First, when governments can choose policy levels optimally they may choose a PTA that they would otherwise have not. Second, the

^{xxx} The political economy of protection is reviewed by [McLaren \(2016\)](#) and [Grossman \(2016\)](#).

preferential tariff itself is now an endogenous object that can be used empirically to learn about the determinants of a PTA such as the degree of import market power (Section 5.2).

If in addition we relax the social welfare assumption, R1, then the government may have redistributive or other political economy motives for a PTA, which generates new predictions. First, political determinants such as the factor endowments of the median voter may become relevant (cf. [Levy, 1997](#)). Second, economic determinants may have a different interpretation. For example, a government that values redistribution to its exporters may prefer partners with high external tariffs because this provides additional protection to its export industry. These agreements are more likely to be trade diverting and thus without R1 governments may support such agreements while the opposite is predicted by the traditional view (cf. [Grossman and Helpman, 1995](#); [Krishna, 1998](#)).

Relaxing R1 and R3 also opens up the possibility that PTAs have a commitment value. This value may be due to the ability of the government to commit to liberalization and thus solve a time inconsistency problem relative to its import competing industry ([Staiger and Tabellini, 1987](#); [Maggi and Rodriguez-Clare, 1998](#)). If we relax R2 to allow for other policies, eg, NTBs, then the PTA also provides a way for the government to commit to its optimal redistribution policy when bargaining with lobbies, as shown in [Limão and Tovar \(2011\)](#).^{sss}

5.1.2.2 Deeper Trade Policy Cooperation and Bargaining Externalities

Deeper trade policy cooperation is important in explaining the trade effects of PTAs. We provided some evidence for this in [Section 4](#) for NTBs and TPU (so relaxing R2 and R5, respectively). These other policy dimensions may generate additional trade for members, but also additional diversion from nonmembers. There is some evidence that PTAs affect both NTBs and TPU and it would be interesting to systematically test if the propensity of certain countries for using this type of policies affects the probability of PTA formation.

Recent work on endogenous trade policy in the context of trade agreements incorporates policy uncertainty caused by either political shocks (cf. [Amador and Bagwell, 2013](#); [Beshkar and Bond, 2015](#)) or political and economic shocks ([Limão and Maggi, 2015](#)). It would be interesting to apply these to a setting with multiple countries where PTAs are explicitly modeled.

Some of the gains from a PTA, eg, commitment and reduction in TPU, would seem to be achievable through multilateral agreements as well. However, countries increasingly pursue them via PTAs. Perhaps there are additional considerations that favor PTAs

^{sss} More broadly, the PTA can generate gains for a set of domestic agents that allows the government to commit reforms that would otherwise be blocked (cf. [Fernandez and Portes, 1998](#)).

that the traditional view ignores such as bargaining and enforcement (R6). Let us consider bargaining first. Customs Unions set a common external tariff so groups such as Mercosur can both explore their import market power (Olarreaga et al., 1999) and negotiate better terms with the rest of the world than any of its individual members. Moreover, in a world where other countries form PTAs it may be optimal for certain countries to respond by doing the same. In Section 5.2, we discuss how this interdependence affects the empirical approach to PTA formation.

Another bargaining advantage in PTAs is that it is not subject to the MFN free riding problems that plague multilateral negotiations (Ludema and Mayda, 2009). Deeper trade policy cooperation may also be easier to enforce in PTAs because (i) there is higher incentive to verify (less free riding) and (ii) there are potentially more policies available to enforce cooperation.^{ttt}

5.1.2.3 Broader Economic Motives

Many PTAs aim to promote investment, not just domestic but increasingly foreign. About 58% of the agreements discussed in Section 2 included clauses for the liberalization of capital movement and 45% included requirements for local content and export performance of FDI. Moreover, there has also been an increase in the number of bilateral investment treaties.

A PTA may increase FDI across members, as they take advantage of vertical specialization possibilities within the PTA. It may also generate FDI from nonmembers to serve the integrated market, ie, export-platform FDI. However, some investment benefits may be offset by local content requirements, which limit the degree of vertical specialization that firms can explore relative to nonmembers.

Is there evidence that PTAs increase FDI? Blomstrom and Kokko (1997) provide case studies and argue there was little FDI increase for Canada during CUSFTA but a significant amount for Mexico from non-NAFTA countries and for Argentina and Brazil in Mercosur. Levy-Yeyati et al. (2003) apply a gravity-type approach to bilateral outward FDI from the OECD to 60 countries in 1982–99. They find that PTA membership increases FDI by about 27% and generates some FDI diversion. Other work analyzing the impact of PTAs on aggregate FDI includes Baltagi et al. (2008).

Recent work employs firm-level data. Chen (2009) applies a gravity-type approach to US multinational sales to its affiliates in manufacturing. After taking PTA endogeneity into account she finds a rise in export-platform FDI but not in FDI between members. Osnago et al. (2015) use firm data and find that vertical FDI increases with the depth of

^{ttt} On the other hand, Maggi (1999) provides a model where multilateral institutions have an advantage over bilateral ones by verifying violations and informing third parties. On the role of policy linkage in enforcement, see Limão (2005) and Maggi (2016).

PTAs. Depth includes the capital movement provisions using the data described in [Section 2](#), so this paper provides a good example of how the rich variation across PTAs can be used. [Tintelnot \(2015\)](#) also explores firm data to quantify the role of multinationals in transmitting technological improvements across countries and the role of trade and investment agreements, eg, between the European Union and Canada, in potentially diverting investment of European Union multinationals from the United States to Canada.

These estimates indicate the importance of relaxing R4 to evaluate the welfare effects of PTAs in settings with FDI and trade in intermediates. In [Section 5.2](#), we discuss the (yet scarce) evidence of the direct effect of FDI on the probability of bilateral preferences and whether their value across industries depends on the fraction of a country's imports produced using FDI owned by its nationals ([Blanchard, 2007](#)).

Certain PTAs explicitly aim to increase productivity and innovation. This objective can be directly seen in the data in [Section 2](#) where 43% of the PTAs include at least some innovation and diffusion provision that promotes technology transfer; joint research projects; exchange of researchers and development of public–private partnerships. It would be interesting to test if PTAs with such provisions do lower information and technology diffusion costs and generate higher rates of innovation. Over 60% of the PTAs also include intellectual property right protection clauses. These can increase innovation incentives and spur additional innovation for at least some the PTA members, but not necessarily all, and may depend on the relative level of development of the members (see [Saggi, 2016](#)).

There are other mechanisms through which PTAs may increase productivity. First, FDI may increase technology transfer. Second, through economies of scale and/or a reallocation of production toward more productive firms (cf. [Head and Ries, 1999](#), for Canada). Third, by generating incentives for exporting firm investments that may increase plant productivity ([Trefler, 2004](#)) and innovation (cf. [Lileeva and Trefler \(2010\)](#) for Canada and [Bustos \(2011\)](#) for Argentina due to Mercosur).

In sum, PTAs reduce trade costs for goods and include provisions for broader economic cooperation that can also lower the costs for FDI and innovation. To understand if these cost reductions and the change in investment incentives in PTAs translate into important innovation and productivity gains we need additional research, which could follow the recent firm-level work cited in this section.

Future work should also examine if services trade expands when related provisions are included in PTAs and if so then whether determinants of services trade increase the likelihood of these provisions.

As noted in [Section 2](#) PTAs also include labor-related provisions. There is interesting work examining the wage effects of certain PTAs (cf. [Hakobyan and McLaren, 2010](#); [Trefler, 2004](#)). Much more could be done to explore the impact of PTAs not just on wages and employment but also on whether relevant provisions have any impact on the outcomes they target such as labor standards.

5.1.2.4 Nonpecuniary International Externalities

There is a long history of preferential trade integration (cf. [Machlup, 1977](#)) and in several important instances it was succeeded by economic and political unions, eg, the Zollverein between German states, the custom union between Italian states. In agreements such as the European Community there is an explicit objective to reduce the probability of conflict via economic integration and accession is conditional on democratization and cooperation in many issues with nonpecuniary externalities.^{uuu}

More broadly, the European Union and the United States commonly provide trade preferences in exchange for cooperation in issues such as the environment, human rights, illicit drugs, and terrorism.^{vvv} As we discussed in [Section 2](#) cooperation in some of these issues is not legally enforceable, but in several cases it does carry a cost. Failure to comply has led certain GSP countries to lose preferential access to the United States. At one point the European Union implemented “special incentive arrangements” whereby developing countries could apply for additional preferential tariff reduction if they satisfied certain labor, environmental, or drug combat criteria.

There is some evidence of the impact of PTAs on cooperation on a subset of the issues discussed earlier. [Mansfield and Milner \(2012\)](#) argue that domestic politics is a critical factor in the decision to join PTAs and democracies are more likely to join. [Liu and Ornelas \(2014\)](#) provide evidence that PTAs increase the probability of survival for democracies. [Hafner-Burton \(2013\)](#) finds human rights improvements in countries with which the United States and the European Union have PTAs with relevant clauses.

[Martin et al. \(2008\)](#) find that multilateral trade openness increases the probability of conflict by decreasing bilateral trade and thus the cost of a bilateral conflict. This suggests a positive role for PTAs in reducing conflict. [Vicard \(2012\)](#) estimates that CU and CM reduce the probability of war between members but shallower PTAs do not. Both he and [Martin et al. \(2012\)](#) find that countries with higher frequency of past wars are more likely to sign PTAs.

The increasing number of PTAs with noneconomic provisions and the rising evidence for their impacts suggests it is important to incorporate them in standard models of policy determination. This would seem to be difficult given the diversity of the issues considered above until we note that several have one common feature: a nonpecuniary international externality. [Limão \(2007\)](#) incorporates that feature in a trade model to derive the incentives for PTAs. He shows that nonpecuniary externalities may facilitate the formation of PTAs for two reasons. First, it increases the set of issues over which they can bargain, which is particularly important between countries of asymmetric economic

^{uuu} The preamble of the European Community explicitly aims for “an ever closer union” and aims to “to preserve and strengthen peace and liberty” and “to strengthen the unity of their economies and to ensure their harmonious development by reducing the differences existing between the various regions.”

^{vvv} On the importance of the exchange of trade preferences for cooperation in nontrade issues, see also [Jackson \(1997, p. 160\)](#), [Abrego et al. \(2001\)](#), and [World Bank \(2000\)](#).

size. Second, this issue linkage improves the ability to enforce cooperation. The model provides predictions for the formation of agreements and the interaction of preferential and multilateral policies. For example, to the extent that these nonpecuniary externalities decay with distance the model provides a nontrade-related explanation for why many PTAs are regional.

5.2 PTA and Preference Determinants

Most empirical work on the economic determinants of bilateral PTAs focuses on trade-related motives, typically trade creation and diversion effects highlighted by the traditional PTA literature. First, we focus on explaining the determinants of the choices made by pairs of countries to form PTAs. We do so both in a setting where the potential agreements are taken as given and independent from each other and then in a setting when some types of interdependencies are controlled for. We argue that interdependence generates strategic incentives that are important in shaping the equilibrium network of agreements and pose a challenge for simple choice-based estimation methods. Second, we consider how richer settings with endogenous tariff levels provide additional determinants and an alternative to the choice-based approach to understanding the motives for the formation of PTAs.

5.2.1 Economic Determinants of PTAs Under Exogenous Tariffs: Binary Choice Approach

5.2.1.1 Independent Bilateral PTAs

Sovereign countries can choose whether to accept a particular PTA depending on whether the payoff is higher than under some feasible alternative. If we restrict the choice to be binary: this PTA or not, then a reasonable criteria for whether x and m implement a PTA is that the resulting change in the government objective is positive for both, ie, $\hat{G}_i > 0$ for $i=x, m$. That is the basic approach followed by [Baier and Bergstrand \(2004\)](#) to motivate a probit estimation of the economic determinants of a traditional PTA with exogenous tariffs. More specifically they use

$$PTA_{x,m} = \begin{cases} 1 & \text{If } \hat{G}_{x,m} \equiv \min(\hat{G}_x, \hat{G}_m) > 0. \\ 0 & \text{otherwise} \end{cases}. \quad (12)$$

They then rely on a specific economic structure to generate hypotheses and guide the choice of economic determinants of PTAs in a cross section. They build on [Frankel et al. \(1995\)](#), who in turn extend the monopolistic competition model that [Krugman \(1991\)](#) used to argue for the natural trading partner hypothesis. Using a simulated solution of the model with exogenous tariffs of 30% on nonmembers they derive larger welfare benefits of an exogenous PTA that eliminates that tariff between pairs of countries that are (i) closer, jointly larger and similar (in terms of GDP) and different in terms of relative endowments and (ii) farther, relatively larger, and similar in terms of relative endowments

relative to the rest of the world. They interpret the factors in (i) as promoting trade creation and those in (ii) as minimizing diversion. Using a probit for a 1996 cross section of PTAs between pairs of 53 countries they find the sign of these determinants conform to the predictions in (i) and (ii).

There are three important points to note. First, this parsimonious set of economic determinants can correctly predict 85% of the 286 PTAs analyzed; this and the consistent sign predictions could be interpreted as strong support for the underlying model. However, a large part of the variation is explained by distance from each other and the rest of the world, which would be a basic prediction from a large set of models where PTAs address a trade externality *or* other cross-border externalities. In fact, [Magee \(2003\)](#) finds that the probability of a PTA in 1998 is higher between closer countries even after controlling for bilateral trade (which he instruments). Second, the joint GDP of the PTA partners is correlated with a host of other possible determinants so we should not make any causal inference, even if the explanatory variable is lagged. Third, any interpretation and causal inference is further complicated by the absence of other controls or alternative approaches that recognize the interdependence of agreements.

5.2.1.2 Interdependent Bilateral PTAs

The formation of any given PTA depends on other existing and potential PTAs in at least two ways. First, holding other agreements fixed, the welfare impact of a PTA depends on the trade that the two members have with nonmembers and thus on any existing PTAs the nonmembers have between themselves or with either of the members. Second, if externalities across PTAs are present and multiple countries can choose multiple agreements then there is a strategic element that affects formation.

How do existing PTAs affect the net benefits for any single new agreement? The marginal benefit for x from lower protection in a PTA market m depends on its exports to m , which in turn depends on how much protection its competitors in that market face and thus on any other PTA that m has. Similarly, the marginal cost to x from deviating from its bilateral optimal policy and lowering protection against m depends on whether x has PTAs with other countries. Finally, all of x and m trade flows also depend on whether other countries in the world have PTAs.

To translate the potential interdependence into testable hypothesis we need to model specific channels. Alternatively, we can ask how the bilateral determinants of traditional PTAs are affected by controlling for interdependence, which is the main question addressed by [Egger and Larch \(2008\)](#). They extend the empirical choice model of [Baier and Bergstrand \(2004\)](#) to a panel setting and include more countries but find similar results for the individual determinants and prediction success rate. They then control for interdependence by including a bilateral distance (or trade) weighted measure of preexisting PTAs for each country pair and find it is positively correlated with the formation and enlargement of PTAs.

The impact of bilateral economic determinants is not sensitive to the inclusion of the interdependence control in [Egger and Larch \(2008\)](#). However, I would argue the existing work does not yet fully establish causal effects due to the potential for endogeneity both for those economic determinants, eg, GDP, and now for the preexisting PTA variable. The latter is assumed to be exogenous but we would expect that the determinants for formation of bilateral PTAs in a region to be persistent and several of those determinants are omitted, so there is potentially a form of lagged-dependent variable bias.^{www}

What specific channels may explain the positive correlation between PTA formation or enlargement and preexisting agreements? [Baldwin \(1995\)](#) proposes a domino theory of regionalism whereby an exogenous integration between w and m diverts trade from x and increases its incentive to form a PTA. Country x 's incentive to join is not sufficient to predict an enlargement—that will also require existing members' approval and thus gain from expansion. But in certain models an exogenous PTA reduces nonmember exports and thus welfare and a further expansion leaves the original members better off, eg, by increasing the market power of a CU relative to the rest of the world and allowing for an endogenous tariff (cf. [Bond and Syropoulos, 1996](#); [Bagwell and Staiger, 1997a](#)).

As discussed in [Section 5.1](#), if a PTA generates trade diversion from nonmembers then the latter are likely to be hurt. [Sapir \(1997\)](#) provides evidence that trade diversion generated by the European Union is associated with subsequent enlargements. [Baldwin and Rieder \(2007\)](#) follow a similar approach where diversion is inferred from a first stage gravity estimation and then used as a determinant of the probability of entering the European Union. [Baldwin and Jaimovich \(2012\)](#) argue this earlier study is “plagued by the endogeneity of the membership” and thus they follow the basic panel approach in [Egger and Larch \(2008\)](#) with a different weighting matrix. The weights are motivated by a contagion model where a PTA between m and w diverts trade from x and this increases the incentive of x to join a PTA with m . The measure suggested by their model is $Cont_{xm} = \sum_w \frac{T_{xm}}{T_x} \frac{T_{wm}}{T_m} PTA_{mw}$. Contagion from m to x is increasing in the importance of the market access to m (measured by its share of x exports) and the market access of its competitor, measured by m 's import share from w . Their baseline estimates use the distance-based weight of [Egger and Larch \(2008\)](#) to control for interdependence and find a positive result, but one that is weaker than that found in the earlier study, perhaps because of the inclusion of a richer set of covariates—including bilateral trade, which is highly significant, and the fraction of exports of a country to PTA partners. When they augment this specification with their contagion-based weight they find only the latter measure of interdependence is significant and conclude this provides support for the contagion/domino hypothesis.

^{www} One way the authors address the potential endogeneity from the interdependence variable is to focus on a cross section and explore the spatial structure of the model to solve for the PTA as a function of economic determinants.

Using a theoretical model to motivate how interdependence works is a clear step in the right direction. However, the interpretation of the interdependence coefficient in [Baldwin and Jaimovich \(2012\)](#) is complicated by the fact that it is obtained after conditioning on many variables—it is not clear if after conditioning on them the model predicts any additional effect for contagion. Moreover, those controls include endogenous variables such as bilateral trade and GDP growth. Future work can provide specifications that follow the theory even more closely and provide a careful identification strategy to address endogeneity.

In the context of interdependent PTAs there is a deeper empirical concern with the studies described: whether the bilateral choice-based estimation is valid. When we expand the set of possible agreements to include more than two countries there are both additional strategic factors in the bilateral decision and a much larger choice set to consider. We discuss this issue further in [Section 6](#). Fortunately, if we consider endogenous policy settings then there are other approaches to learning about the determinants of PTAs. We now examine these alternatives.

5.2.2 Alternative Determinants and Approaches to PTA Formation Under Endogenous Tariffs

As described in [Section 5.1](#), the traditional PTA literature treats the initial and final tariffs as given and this is the approach of most of the choice-based empirical PTA formation literature discussed earlier. Allowing for endogenous policy levels provides alternative interpretations to existing findings and a richer set of determinants. Moreover, it allows for additional tests of the determinants of PTAs.

5.2.2.1 Alternative Determinants and Interpretations

If we allow governments to optimally choose policy levels then the degree of change in the bilateral tariffs due to a PTA reflects the incentives for the initial and final tariffs. In theory, WTO members are subject to certain constraints on preferential tariffs, but there are numerous exceptions that countries can and do build into PTAs. Even if the WTO enforced the constraint that PTAs under Article XXIV set most preferential tariffs at zero, the initial tariff could still reflect terms-of-trade and political economy motivations. We consider each in turn.

[Magee \(2003\)](#) empirically examines some political economy determinants of PTA formation. For example, in [Levy \(1997\)](#), the median voter decides whether a PTA is formed, this model predicts that agreements are more likely if partners have similar capital-labor endowments, which is the opposite of the prediction tested by [Baier and Bergstrand \(2004\)](#) using welfare maximizing models. [Mitra et al. \(2002\)](#) provide evidence based on [Grossman and Helpman \(1994\)](#) that indicates democracies place higher weight on social welfare than dictatorships. This suggests that welfare-enhancing agreements are more likely to be reached if both countries are democracies. Magee finds supporting

evidence for both of these determinants of PTAs using a probit in a panel setting. He conditions on bilateral trade, which is instrumented, and even then some of the distance and economic variables—meant to capture trade effects of PTAs in earlier work—remain significant, and some variables such as the capital–labor ratio, switch signs. This suggests caution in interpreting the effects of those variables in [Baier and Bergstrand \(2004\)](#) as purely economic determinants of PTAs signed by welfare maximizing governments.

The bilateral terms-of-trade motive for PTAs has not been explicitly tested. One way to do so would be to use the country and product characteristics that [Broda et al. \(2008\)](#) identify as determinants of import market power. These include a large import share of a good relative to the nonmembers and the degree of product differentiation. Another determinant of import market power is remoteness from other markets. There is evidence of remoteness predicting PTA formation, which has been interpreted as capturing smaller trade diversion potential. An alternative interpretation is that remoteness confers regional market power and so an incentive for a PTA. Instead of using country characteristics that may capture other factors, we could directly use the country–industry measures of market power estimated by [Broda et al. \(2008\)](#) and ask if PTAs are more likely between country pairs with larger share of trade in those industries and if preferential tariff cuts are larger.

It would also be interesting to understand why countries choose agreements with specific types of policies, eg, investment provisions. Only recently has this question started to be addressed. [Orefice and Rocha \(2014\)](#) find that the probability of formation of deeper PTAs is higher between countries with a larger share of trade in intermediate inputs, particularly if one country is low income. Thus the potential to share production networks appears to be one motive in the selection of PTA partners.

5.2.2.2 Alternative Approaches: Determinants of Votes and Policy Outcomes

Endogenous tariff setting also opens up a broader set of approaches to study the determinants of PTAs. First, we can consider the determinants of outcomes of votes. [Baldwin and Magee \(2000\)](#) find that campaign contributions influenced US legislator votes on NAFTA and that certain economic conditions in each member’s district affected the vote outcome. [Conconi et al. \(2014\)](#) find that congressional term length and election proximity affect the likelihood of support for trade liberalization bills such as NAFTA. [Beaulieu \(2002\)](#) finds that both industry of employment and type of factor (skilled/unskilled) of individual Canadian voters were significant determinants of their vote on whether to support that country’s PTA with the United States.

Another approach is to analyze the determinants of preferential tariff rates, which can vary bilaterally and at the product level. We highlighted the bilateral terms-of-trade externality as a potential motive for PTAs. While recent work has established the importance of this externality for unilateral and multilateral tariffs much less is known about its role in shaping bilateral tariffs such as those governed by PTAs. One exception is

Olarreaga et al. (1999) who find that Mercosur's common external tariff is higher in products where it had more import market power and this measure (Mercosur share of world imports in each product) can explain up to 28% of the tariff variation. This indicates the potential terms-of-trade gain from a CU derived from setting tariffs jointly.

More recently, Blanchard et al. (2016) provide indirect evidence of the TOT motive for PTAs. They introduce supply chain linkages into an endogenous trade policy model with TOT externalities. The model predicts that the noncooperative bilateral tariff decreases with the share of domestic value added in imports. They find support for the prediction in countries without PTAs, by using bilateral product tariffs in 14 economies in 1995–2009. If PTAs internalize the TOT externality then preferential tariff variation within an agreement should not vary with domestic value added, and that is their finding. This paper provides a good example of the type of detailed evidence guided by theory that addresses identification concerns and can be used to learn about motives for the formation and shape of policy in PTAs, particularly the role of supply chains.

In Section 4, we discussed how PTAs could affect the degree of vertical specialization. Blanchard (2007) provides a theory where a country's protection declines with the fraction of imports that are produced using FDI made by its nationals. If a US tariff reduces the price received by a foreign exporter then some of that cost translates into lower profits for any US multinational involved in the production of that good abroad. This basic insight suggests an incentive for lower tariffs on partners where that FDI is most prevalent and thus for PTAs. The panel evidence in Blanchard and Matschke (2015) for the United States between 1997 and 2006 supports this prediction. They find that a 10% increase in US multinational exports to the United States reduced its preferential tariff by about 4 percentage points.

Another well-documented determinant of unilateral trade policy is the role of lobbies. Less is known in the context of PTAs even though the additional variation in preferential tariffs may be very informative. Kee et al. (2007) find that foreign lobbies' contributions to the United States affect the latter's preferential tariffs to countries in the Americas.

5.2.3 Summary

In sum, there has been some advance in empirically identifying the basic trade-related mechanisms that underlie the formation of traditional PTAs: creation, diversion, and price effects. The research focusing on the determinants of PTAs has confirmed the importance of bilateral trade and thus the importance of the endogeneity concerns in gravity estimates. The evidence for other determinants of trade creation/diversion in the formation of PTAs is suggestive but more work is required to establish causal relationships. The same is true for the role of past PTAs by a country and its partners. A similar choice-based estimation could be used to explain the determinants of the *type* of agreement. However, in the presence of interdependence of PTAs and multiple choices an alternative approach may be required.

One promising avenue to understand the determinants of PTA formation and the depth of cooperation is to explore preferential tariffs and other product level data. This may allow us to test sharper predictions, establish causal effects and identify certain structural parameters that may be used to quantify interesting counterfactuals. For example, whether uncertainty regarding multilateral tariffs or temporary preferences, such as those developing countries receive via the GSP, makes it more likely to seek a preference bound at zero.

6. AGREEMENT AND POLICY INTERDEPENDENCE

In the previous section, we noted that interdependence between PTAs poses a challenge to identifying their bilateral determinants using a choice-based approach. We then argued that exploring the determinants of endogenous preferential tariff levels within any given agreement could provide additional insights about the incentives for their formation and policy depth. We now provide a brief discussion of the basic issues when agreement decisions are interdependent and then turn to evidence that examines the interaction between bilateral policy levels covered in a PTA and those not covered by that agreement.

6.1 Agreement Interdependence: Approaches and Open Questions

In the context of interdependent PTAs there is a deeper empirical concern with some of the studies in [Section 5.2](#): whether the bilateral choice-based estimation is valid. When we expand the set of possible agreements to include more than two countries there are both additional strategic factors in the bilateral decision and a much larger choice set to consider. We illustrate the issue and note how the theoretical approaches addressing the equilibrium structure of agreements require us to consider additional empirical approaches to study the formation of PTAs and the interdependence of policies.

To illustrate the basic issue consider an example with four countries in two blocs, x and m ; w and w' . Suppose that w and w' have a cost of forming a PTA that is sufficiently low between themselves that they always form it but prohibitively high with either x or m . If x and m are deciding on PTA between themselves after w and w' have already formed one then the decision between x and m may be represented by the bilateral choice Eq. (12).^{xxx} The empirical work in [Section 5.2](#) that controls for interdependence can address this specific case. But suppose now that each of the two blocs of countries is simultaneously deciding between three possible outcomes: cooperation in (i) no policy, (ii) bilateral tariff with a country in its bloc, or (iii) bilateral tariff with a country in its bloc and multilateral tariff with the remaining. There are now nine possible outcomes to consider and this number increases exponentially in the number of such blocs. The

^{xxx} This presumes the costs of dissolving the PTA between w and w' are high so it will persist regardless of others.

choice set is even larger if we reduced the costs of PTA formation across blocs and allow for overlapping membership in different PTAs.

The interdependence between agreements raises three basic questions: what are the possible PTA network outcomes; what theoretical criteria should we use to predict which will emerge; and how might the network evolve. In settings with transfers between countries we have strong results on the existence of necessarily welfare-enhancing CU (Kemp and Wan, 1976) and FTA (Panagariya and Krishna, 2002), which leave nonmembers indifferent. One implication of this literature is that there may exist a PTA path for welfare maximizing governments to expand a PTA until it includes all countries. But transfers may not be available, governments are not welfare maximizers and even if they were the strategic interaction between them they may generate suboptimal outcomes from the global perspective.

At a broad level, there have been two theoretical approaches to address endogenous formation in the presence of this type of bargaining externalities across PTAs. One approach restricts the set of possible outcomes by considering a small number of countries and then considers alternative methods to select the more “likely” outcome by determining which outcomes are in the core (Riezman, 1985) or are the equilibrium of some game that can involve either sequential bargaining (Aghion et al., 2007) or the Nash equilibria of a simultaneous game (cf. Saggi and Yildiz, 2010; Saggi et al., 2013).^{yyy} The simultaneous game approach typically requires some mechanism such as coalition proofness to select between multiple Nash equilibria. The multiplicity can arise because countries may prefer no agreement if no others have them but are better off in a bloc if other blocs form. The second approach allows for a larger number of countries and uses network theory to examine which bilateral networks are stable (Goyal and Joshi, 2006; Furusawa and Konishi, 2007). The larger number of countries comes at the cost of focusing on pairwise stability, ie, whether to form or sever a single PTA, holding all else constant.

The theoretical approaches addressing endogenous formation provide a number of interesting insights. The central implication I want to draw for the current discussion is that the choice set in the bilateral criterion in (12) is restrictive. If the number of potential choices was sufficiently small we could consider nested or multinomial approaches (eg, for each pair model whether to have an agreement and if so how many of which type). This type of approach is followed by Egger et al. (2013) to examine the determinants of the bilateral choice between mutually nonexclusive types of agreements (on goods, services, taxation, investment, or currency unions) and the resulting impact on outcomes such as trade and FDI. They find that agreements on goods have a larger trade effect when combined with one on investment and the same is true more generally of combining different types of agreements on any given outcome.

^{yyy} Yi (1996) uses a noncooperative approach for many symmetric countries.

In general, even multinomial approaches will be problematic given the large number of choices (eg, partners) and their interdependence. For example, in 2010 a bilateral country pair with a PTA had on average 52 other bilateral links. It would be interesting to consider empirical approaches to related choice problems used in other fields, such as the decisions of firms to enter multiple markets in the presence of spatial competition (cf. [Aguirregabiria and Suzuki, 2016](#)).

Fortunately, if we consider endogenous policy settings then there are approaches that do not rely on choice-based estimation to learn about how agreements interact, eg, examining how preferential policies affect the incentives for policy against nonmembers, which we examine next.

The open questions that remain are not simply empirical. In [Section 2](#) and its Online Appendix, we noted the parallel between the evolution of cooperation within the WTO and the type of provisions that its members now include in PTAs; we also noted that the latter go farther than the WTO in several respects. This raises the question of whether and why there are limits to the depth and breadth of cooperation in the WTO. If so, then is this because of higher heterogeneity in preferences, enforcement problems, free riding, and the regional scope of externalities? Is it related to the more open membership that characterizes the WTO? To properly address these questions we need a model of evolution of cooperation within agreements that rationalizes what we observed in the GATT/WTO. This may point to any potential cooperation constraints the WTO and provide insights on why its members increasingly pursue cooperation via PTAs.

6.2 Policy Interdependence

As we discussed in the introduction, PTAs have proliferated at the same time that WTO membership expanded and MFN tariffs were declining. A large share of trade between WTO members is carried out by groups that have bilateral PTAs. Thus we now examine the potential for policy interdependence. Theory can provide some guidance regarding the set of policies that are likely to be interdependent, eg, same industry tariffs or NTBs against different partners, but predictions regarding the direction of causality and the sign, eg, do preferential tariffs raise or lower the multilateral tariff or vice versa, are more sensitive to specific modeling assumptions and thus they are ultimately empirical questions. The clear potential for two-way causality implies that identification issues are paramount and I discuss these as well in the context of existing estimates.

What do we mean by policy interdependence? Heuristically, we say that a PTA and a non-PTA policy are interdependent if changes in the value of one affects the incentives for the equilibrium value of the other. So naturally we are considering an endogenous policy setting. We distinguish between two types of interdependence: multilateral, eg, between preferential and MFN tariffs for a country, and bilateral, eg, between a cooperative tariff set between PTA partners and noncooperative NTBs between them. Both types can affect the decision to form agreements.

6.2.1 Multilateral Tariff Interdependence

More research has focused on multilateral interdependence because it can also directly affect the protection faced by nonmembers. This research was motivated by the debate on the effects of PTAs on multilateral liberalization triggered by the difficulties in completing multilateral trade negotiations in the Uruguay Round. This generated a voluminous theoretical literature with arguments supporting the notion that PTAs are, as [Bhagwati \(1991\)](#) put it, a stumbling block to multilateral trade liberalization (MTL) and others against it. We focus on the more recent work that provides empirical evidence relevant to this question.

Multilateral tariff interdependence refers to the relationship between preferential tariffs and nonpreferential ones that are applied to nonmembers. There is potentially two-way causality since the optimal preferential tariff depends on the value of the tariff on nonmembers and vice versa as illustrated by the following simple example. Suppose the initial multilateral tariff of a country x is zero, $\tau_{x,w} = 0$. If there is a constraint that the multilateral tariff that x sets on w cannot be changed as a result of PTAs, ie, $\hat{\tau}_{x,w} = 0$, then x is unable to offer any preferential tariff to m (unless it is a subsidy) so there would be no market access value to this PTA for m . This example also illustrates how preferential tariffs can affect the incentives for multilateral liberalization. If there was no multilateral constraint and the PTA was sufficiently valuable to the members they would have an incentive so increase multilateral tariffs, ie, set $\hat{\tau}_{x,w} > 0$.

The empirical analysis has focused on the effect going from preferential to multilateral tariffs with the goal of establishing whether $\hat{\tau}_{x,w} \neq 0$. We are also interested in the sign of this relationship: if $\hat{\tau}_{x,w} > 0$ then the PTA may impose an additional cost of nonmembers in the form of lower export prices. There is an additional source of interdependence relevant for the stumbling blocs question. Namely whether a nonmember, w , may respond by changing its tariffs on a country x when the latter forms a PTA with m . This tariff response is harder to identify empirically, but it is more likely to be an increase in protection if $\hat{\tau}_{x,w} > 0$, a reciprocal response. So evidence for whether tariffs are raised against nonmembers is important both through its direct and reciprocity effects on multilateral liberalization.

6.2.1.1 Basic Hypotheses

Some theories predict $\hat{\tau}_{x,w} > 0$ and others the opposite, as we may expect in a second best setting. Our objective is not to lay out all possible theories but rather to ask what are some relevant general conditions leading to either case and what insights there are for the estimation approach.^{zzz}

To gain some insight about the determinants of tariff interdependence and the implications for empirical testing we consider a simple setting. Suppose the government has an

^{zzz} See [Bagwell and Staiger \(1998\)](#), the contributions in [Bhagwati et al. \(1999\)](#) and the surveys by [Winters \(1999\)](#), [Freund and Ornelas \(2010\)](#), and discussion in [Bagwell et al. \(forthcoming\)](#).

objective, $G_x(\tau_{xw}, \tau_{xm}, \cdot)$, that is a function of its tariff on the PTA partner (τ_{xm}) the non-member (τ_{xw}) and the tariffs it faces (omitted). If the government could unilaterally choose τ_{xw} after setting its preferential tariff then it would satisfy a first-order condition, $\frac{\partial G_x(\tau_{xw}^N, \tau_{xm}, \cdot)}{\partial \tau_{x,w}} = 0$. So a PTA, which leads to a lower τ_{xm} causes $\hat{\tau}_{x,w} > 0$ if the tariffs are strategic substitutes in G_x or the opposite if they are complements. This condition depends on the political and economic structure and so it is model specific. In discussing the empirical results that find a positive or negative relationship between tariffs we refer to specific theoretical models they are consistent with. But we offer two observations that should be independent of that sign. First, strategic interdependence is more likely between tariffs in similar goods, denoted by $k \in K$, which suggests testing the impact of some measure of $\tau_{x,m,k}$ on $\tau_{x,w,k}$. Second, if k is either not imported from the PTA or does not receive a preference then we expect $\hat{\tau}_{x,w,k} = 0$, or at least smaller than the effect on other goods in the same industry imported preferentially.

When we focus on interdependence at the good and/or industry level I would argue that there is one force pushing toward substitution present across different models that gives rise to what I call the preference erosion hypothesis. The market access gain from a PTA in a given good depends on the preference margin relative to the nonmember, which is simply $\tau_{x,w,k} - \tau_{x,m,k}$. So if for an initial value of $\tau_{x,w,k}$ the optimal preferential tariff is positive there may be no need to increase $\tau_{x,w,k}$ to increase the preference margin. But if a country has already liberalized extensively then the preference margin constraint will bind. In a setting with multilateral negotiations this incentive translates into less multilateral liberalization to avoid preference erosion. The preference erosion concern has often been voiced in the WTO by developing countries afraid to have GSP preferences eroded. Its potential effect on multilateral liberalization was anticipated by opponents of the GSP when it was originally proposed (cf. Johnson, 1967, p. 166). It is important to notice that a similar logic applies to other types of preferences unrelated to GSP, as subsequent evidence finds. The two testable predictions from the preference erosion hypothesis are that $\hat{\tau}_{x,w} > 0$ is more likely for (i) countries/goods with an initially low $\tau_{x,w}$ and (ii) goods with a binding preference margin, ie, when $\tau_{x,m,k} = 0$.^{aaaa}

6.2.1.2 Estimation

To answer whether a country's PTAs affect its nonpreferential liberalization we require a reasonable empirical counterfactual. The basic challenge is that the theoretical object of interest, $\hat{\tau}_{x,w}$, compares protection relative to an unobserved situation. One could compare how changes in the number of a country's PTAs affect its aggregate MTL over time. But, with so many other possible determinants of aggregate MTL changing between trade rounds it would be hard to convincingly attribute any differences in MTL solely

^{aaaa} This effect should be stronger for those goods where the relevant NTBs are highly constrained by the WTO.

to PTAs. A similar concern applies if we compare the aggregate liberalization of countries with and without PTAs, as illustrated by the findings in [Foroutan \(1998\)](#). She finds lower average MFN tariffs for Latin American countries with PTAs after the Uruguay Round but notes that no causality can be drawn from such a correlation because those countries were moving away from import substitution during the 90s, which implied considerable unilateral liberalization independently of any effects from PTAs.

[Limão \(2006\)](#) proposes a difference-in-difference approach where changes in non-preferential tariffs for each good k in the United States are a function of the status of the good changing from non-PTA to PTA between the Tokyo and Uruguay Rounds. By exploring variation in PTA status over time and across goods he controls for any aggregate (and sector) unobservable heterogeneity. To compare this and other approaches it is useful to consider the following basic estimation equation.

$$\Delta\tau_{x,w,k} = \gamma_x PTA_{x,m,k} + X'\beta + u_{x,w,k}, \quad k = 1, \dots, N \quad (13)$$

Using changes in US bound tariffs where k is an HS-8 product [Limão \(2006\)](#) estimates a positive γ_x , so multilateral tariffs for PTA goods increased relative to non-PTA goods. He argues this represents a causal effect once we control for various determinants of the multilateral tariff, such as reciprocity and bargaining power measures, and instrument the PTA variable. A key concern is that countries seek preferences in goods where they expect smaller multilateral reductions. To address this he provides and tests the validity of alternative instruments for $PTA_{x,w,k}$.^{bbbb} He finds that the United States would have cut multilateral tariffs by twice as much in the absence of its PTAs.

[Karacaovali and Limão \(2008\)](#) find similar results for the European Union. They also test and find support for two situations where their model predicts no stumbling block effect: (i) in goods with positive preferential tariffs and (ii) for CU. These are also consistent with a nonbinding preference margin constraint since in (i) the preferential tariff can still be reduced and in (ii) any “missing” margin can be offset via direct transfers.

[Ketterer et al. \(2015\)](#) apply the same approach as [Limão \(2006\)](#) and find that preferences given by Japan lead it to keep its multilateral tariffs 1.7 percentage points higher in the UR. [Ketterer et al. \(2014\)](#) find a positive but insignificant effect of almost all of Canada’s preferences on its multilateral tariffs in the UR. The exception is NAFTA preferences, which they conclude generated a reduction in Canadian multilateral tariffs. They interpret this as evidence for a rent destruction effect, which is predicted by models such as [Ornelas \(2005\)](#). This is one possible interpretation but other models also predict a complementarity of preferential and multilateral tariffs (cf. [Bagwell and Staiger, 1997b](#)). One concern with the estimates for Canada is that none of the specifications control for

^{bbbb} In particular the instrument for whether the good is exported by the PTA partner *and* receives a preference is whether the good is exported before the multilateral tariff changes, which is correlated with the variable but excludes the potentially endogenous components, ie, the preference. The author also uses other instrument such as transport costs and world price variables.

all agreements simultaneously, and, since $PTA_{x,m,k}$, is clearly correlated across agreements, we cannot be certain of what the individual agreement effects are or whether any is significant when all are included.

Ketterer et al. (2014) argue that the difference between their result for Canada and the stumbling bloc findings in earlier work for the United States and the European Union is because these countries have agreements with small nations from whom they extract nontrade benefits. But that is incorrect since the earlier studies find the stumbling bloc effect is present both for GSP recipients, agreements with small developing countries and agreements with developed countries when using specifications that include all of the agreements separately.^{cccc} An alternative explanation for the difference in the results for Canada is that before the UR its multilateral tariff was considerably higher, about 12%, than that of the European Union 8%, the United States, or Japan (about 7%). Recall that one of the predictions from the preference erosion hypothesis was that a stumbling bloc was less likely when the multilateral tariff is higher.

Another strand of the literature has focused on unilateral (rather than negotiated) tariffs on nonmembers. Tovar (2012) uses product data and finds evidence of higher external tariffs for El Salvador, Guatemala, Honduras, and Nicaragua in 2005–09 due to preferences they granted the United States under an FTA. Other studies for Latin America use industry data and find different effects. Bohara et al. (2004) estimate that the Argentine unilateral tariffs were lower in industries with higher imports from MERCOSUR. Estevadeordal et al. (2008) study 10 Latin American countries between 1990 and 2001 and find that countries reduced nonpreferential tariffs by more in industries with larger preference margins. They acknowledge the potential for endogeneity and argue that it is addressed by instrumenting the change in preference margin given by a country with those of its PTA partners. They find no effect for CU.

Crivelli (2014) uses the data and basic approach in Estevadeordal et al. (2008) with two key differences: controlling for unobserved industry shocks and estimating heterogeneous responses. She first replicates their building bloc effect but then shows it becomes insignificant after including industry–year effects, which control for unobserved industry incentives to lower both preferential and nonpreferential tariffs. She then estimates a heterogeneous response and finds that a 1% preference margin is associated with a 0.18% decrease in nonpreferential tariffs for goods with initial MFN tariff higher than average but no effect for the other goods. She concludes this reflects an incentive to minimize the PTA cost of lost tariff revenue from nonmembers first proposed in Richardson (1993).^{dddd}

^{cccc} For example, in Limão (2006) the stumbling bloc effect for NAFTA and GSP is similar, Karacaovali and Limão (2008) find similar effects for GSP as well as EFTA and central European countries.

^{dddd} Pardo et al. (2009) examine tariffs for Asian countries in 1992–2007 and conclude that “preferential liberalization tends to precede external tariff liberalization.”

In sum, there has been considerable progress in the estimation of multilateral interdependence. There is still work to be done to establish causality in some settings and test alternative mechanisms. However, there are two basic robust findings thus far. First, there is a stronger stumbling bloc effect for countries and goods with lower initial multilateral tariffs. Second, there is no interdependence for CU but beyond this there is yet no definitive evidence of different effects across types of agreements.

6.2.1.3 Reciprocity and Welfare Effects

To determine the effect of PTAs on worldwide liberalization we would ideally also like to determine the nonmember tariff changes in reaction to PTAs. This has only been indirectly analyzed by modeling and estimating multilateral reciprocity effects in the framework of Eq. (13). [Limão \(2006\)](#) finds that the US multilateral liberalization was reciprocal so the direct stumbling bloc effect estimated for the European Union in [Karacaovali and Limão \(2008\)](#) would have led to smaller reciprocal reductions by the United States and vice versa.

It would also be interesting to understand the welfare impacts of multilateral interdependence. [Limão and Olarreaga \(2006\)](#) argue that import subsidies can eliminate the stumbling block effect of PTAs that arises from concerns with preference erosion.^{eeee} They then use the estimates in [Limão \(2006\)](#) to compute the counterfactual welfare gains through the additional MTL resulting from switching away from preferences to this subsidy scheme. They find nonnegligible gains for countries that grant preferences, those receiving them and outsiders even if the switch occurs only for the preferences given to least developed countries by the United States, European Union, and Japan.

6.2.2 Additional Dimensions of Interdependence

6.2.2.1 NTBs Against Nonmembers

PTAs can also change the incentive for NTBs against nonmembers. It is important to understand if they do so since NTBs account for an increasing share of protection. Moreover, Article XXIV does not prevent an increase in NTBs after PTAs, so it may be easier for countries to explore them, and thus for us to detect these effects of PTAs. [Bown et al. \(2014\)](#) discuss some case studies and the challenges to analyzing this question more systematically.

[Prusa and Teh \(2010\)](#) find that PTAs that include antidumping provisions reduce filings between members by 33–55% but increase them against nonmembers by 10–30%. [Limão and Tovar \(2011\)](#) find that the external tariff constraints that Turkey had to accept when it formed a CU with the European Union increased the probability and AVE of Turkish NTBs. Further work on this topic is important.

^{eeee} Import subsidies achieve this because, if they are set at fixed rate, they are independent of the MFN tariffs and thus the latter can be reduced without any preference erosion.

6.2.2.2 Received Preferences

To our knowledge, there is no evidence on the effect of new preferences on the external protection of the receiving country. In part this reflects an empirical identification challenge: the strategic interdependence criterion indicated a clear path to relate preferences received in the same good to different partners, which is not obvious for a preference received. However, [Özden and Reinhardt \(2005\)](#) find that the elimination of GSP preferences is correlated with subsequent aggregate multilateral liberalization of that beneficiary country.

6.2.2.3 Effect of Multilateral Tariffs on Preferences

At the start of this subsection we noted that multilateral tariffs can also affect preferential tariffs. There is little evidence about this direction of causality, which is hard to establish. [Fugazza and Robert-Nicoud \(2014\)](#) find that the United States is more likely to extend duty free preferences in goods where it cut its multilateral tariff the most in the UR. It is not clear if this reflects any dynamic change in incentives or simply the desire to maintain a fixed preference margin after the multilateral tariff is reduced.

6.2.2.4 Bilateral Interdependence

What effect do cooperative preferential tariffs between two countries have on other policies between them not covered in the agreement? Understanding this type of interdependence can help explain the policy scope of agreements, eg, if tariffs and certain NTBs are highly substitutable then an agreement that only includes the former would be unraveled by the use of NTBs and thus have little value, which may explain why almost all the agreements in [Table 1](#) include both. It would be interesting to explore that detailed database to determine the extent of substitution between policies included and those excluded from agreements. It would also be interesting to test if such substitution increases the probability of subsequent cooperation in those NTBs.

6.3 Wither Multilateral Tariffs, Wither Diversion, and Nonmember Discrimination?

After the completion of the UR the number of PTAs increased even further and the “new” round started in 2001 is yet to be completed. But according to [Baldwin \(2016, p. 112\)](#) “the global tariff-cutting since the rise of regionalism has proceeded as quickly as ever, but outside the WTO [and] as a result, the specter that regional trading agreements would inefficiently divert trade never really appeared” and notes that the low MFN tariffs and small applied preference margins in PTAs, which we also noted before, imply that “bilateral and regional trade agreements provide a relatively small incentive to divert trade” and points to the building bloc evidence in [Estevadeordal et al. \(2008\)](#) as an additional factor that should ease concerns about diversion.

My assessment of the evidence for the diversion potential of PTAs is less optimistic. First, the evidence that various PTAs hinder tariff liberalization cannot be ignored and we still know little about the impacts on NTBs. Second, the evidence in [Section 3](#) shows that the PTA trade effect on members relative to nonmembers is large and not fully explained by tariffs, so even if applied tariffs are small and generate small incentives for diversion, other barriers may not be. Third, in [Section 4](#) we provided evidence that the trade elasticity is higher when tariffs are low and more certain. So, even if PTAs are simply removing small applied tariffs and uncertainty about them, they can divert large amounts of trade toward members and away from nonmembers. Moreover, the evidence shows the trade effects of PTAs increase over time suggesting there could be investment and growth effects that amplify initially smaller effects. Fourth, there is some evidence of PTA contagion ([Section 5](#)) where a plausible underlying mechanism is trade diversion. In sum, PTAs may have many positive effects but we should continue to investigate whether and how they affect trade (and resulting welfare) of members and nonmembers.^{ff}

In [Section 5](#), we discussed various other motives for pursuing PTAs, so even if these agreements do not divert trade they may divert FDI (as found by [Tintelnot, 2015](#)) and generate a host of externalities for any countries excluded from the negotiations. One potential solution that has been advanced to minimize the effects of PTA on nonmembers is to pursue an open membership policy. In principle this can internalize the effects of a given PTA, but in practice it neglects the fact that the gains or ability to enforce cooperation in certain PTAs may be due to their closed membership. Models of the evolution of cooperation may provide some insight into this issue. It would also be useful to estimate if PTAs affect outcomes for nonmembers beyond trade.

7. LESSONS AND FUTURE RESEARCH

Given the breadth of the chapter and diversity of agreements considered most conclusions are qualified but there are some important lessons and guidance for future research in each section, some of which I highlight here.

7.1 Deeper and Broader Policy Cooperation

From the preamble of a typical recent PTA it is clear they address far more than tariff reductions. This has long been obvious for agreements such as the European Union; [Section 2](#) shows these additional policy dimensions are now widespread and that “deeper” PTAs account for a large and growing trade share of bilateral world trade.

^{ff} Also, the move in PTAs toward addressing “behind-the-border” policies may limit the ability to discriminate against nonmembers in certain policies, as [Baldwin \(2016\)](#) points out, but not in others and these types of policies can have trade impacts that exceed those of small tariff preferences.

Moreover, the evidence in [Section 3](#) shows that PTAs in 1990–2010 had a very small effect on tariffs applied between members, around 2 log points.

These facts warrant a shift in research beyond the traditional view of PTAs as static tariff reductions. Whether and how to do so led us to systematically identify and group 52 policies along two dimensions of cooperation: depth and breadth. Future research can pursue each of these dimensions in turn to determine their importance not just on trade but also on other economic outcomes. One approach is to explore the rich variation in policies across PTAs or the degree of their legal enforceability within PTAs. Another approach is to model specific features and test them using detailed data, as illustrated by the work on policy uncertainty in [Section 4](#). The next challenge is to explain not just the impacts of specific policies in PTAs but also how those policies interact and what determines their inclusion in different agreements.

7.2 Trade-Related Effects on Members

A common important element across diverse PTAs is their aim to increase bilateral market access. Do they achieve this aim? The first clear lesson related to the bilateral trade effects of PTAs is that these are large on average, increasing over time and heterogeneous across agreements. The second lesson is that these effects cannot be fully explained by traditional models where PTAs amount to static preferential tariff reductions. The third lesson is that trade policy cooperation in modern PTAs aims at lowering not simply applied bilateral trade costs, eg, tariffs and NTBs, but also their uncertainty. Recent evidence indicates the uncertainty channel helps to explain large and heterogeneous PTA trade effects.

Future research should explore additional mechanisms that explain the large and heterogeneous PTA trade effects. One useful approach is to model deeper trade policy cooperation in richer economic settings, eg, with investment and intermediates, and estimating the effects on trade and related firm decisions. Doing so can further bridge the current gap between theory and quantitative work, that focuses on tariff changes under a constant trade elasticity, and empirical research that estimates average treatment effects using a PTA dummy.

7.3 PTA Formation and Policies

There has been some advance in empirically identifying the basic trade-related mechanisms that underlie the formation of traditional PTAs: creation, diversion, and price effects. The research focusing on the determinants of PTAs has confirmed the importance of bilateral trade and thus the importance of the endogeneity concerns in gravity estimates. The evidence for other determinants of trade creation/diversion in the formation of PTAs is suggestive but more work is required to establish causal relationships. The same is true for the role of past PTAs by a country and its partners. A similar choice-based

estimation could be used to explain the determinants of the type of agreement. However, in the presence of interdependence of PTAs and multiple choices an alternative approach may be required.

One promising avenue to understand the determinants of PTA formation and the depth of cooperation is to explore preferential tariffs and other product level data. This may allow us to test sharper predictions, establish causal effects, and identify certain structural parameters that may be used to quantify interesting counterfactuals.

There is still scant evidence on the mechanisms underlying nontraditional motives for PTAs. Some of those motives are reflected in the broader dimensions of PTAs, described in [Sections 2 and 5](#). For example, the empirical effect of PTAs on FDI is still mixed but there is some evidence that FDI affects the degree of preferential treatment. We do not yet know if PTAs affect technology transfer or intellectual property significantly. There are still only a few studies on the labor market impacts of PTAs. In terms of the noneconomic dimensions, there is some evidence of the positive effects of PTAs on bilateral conflict but almost none addressing the environment, human rights, or democracy. These are all interesting areas for future work, whether to understand the motives for PTA formation or, more generally, to use these large and frequent shocks as a way to identify the impacts of globalization.

7.4 Interdependence

The other central feature of the current trading system is the interdependence of agreements. The share of WTO country pairs with PTAs was over 25% in 2010 and their corresponding world trade share was 55%. Moreover, in 2010 a bilateral country pair with a PTA had on average 52 other bilateral links. Understanding this interdependence is central to explaining the formation and the worldwide effects of PTAs. There is evidence that past PTAs by a country or its partners are correlated with the probability of current PTA formation. This remains an interesting area to explore but future work should tackle the identification and conceptual challenges inherent in a world where the choice set of partners and policies is so large and there are strategic interactions.

Whether PTAs are a stumbling or building bloc to multilateral liberalization remains an important question. The substantial theoretical literature with arguments in both directions paved the way for empirical evidence. There has been considerable progress estimating the effect of preferential tariffs on multilateral or unilateral tariffs from which I draw two robust findings. First, there is a stronger stumbling bloc effect for countries and goods with lower initial multilateral tariffs. Second, there is no interdependence for Customs Unions. But beyond these there is yet no definitive evidence of different effects across types of agreements. Future work should analyze how PTAs affect the incentives to change deeper policy cooperation, eg, uncertainty and NTBs, toward nonmembers.

Even if PTAs do not substantially divert trade they may divert FDI and generate other externalities toward nonmembers. Modified rules that allow for more open membership can help internalize some effects of a given PTA. But in practice such suggestions neglect that the gains or ability to enforce cooperation in certain PTAs may be due to their closed membership. Models of the evolution of cooperation may provide some insight into why countries increasingly move their cooperation beyond open membership organizations such as the WTO and how any resulting negative externalities on nonmembers may be minimized.

In conclusion, PTAs have been, and are likely to continue being, a key source of trade policy reform and an exciting area for research. Their estimated bilateral trade effects and their policy scope indicate the value of augmenting the policy and economic structure of our models relative to the traditional view of PTAs as static tariff reductions. Doing so is important to improve our understanding of the design of PTAs and their effects. There is also enormous potential for new empirical research. The frequency of new PTAs and the variation in their policies and enforceability across agreements are but a few examples of the shocks that can be explored to examine the impacts of globalization on workers, firms, consumers, and a variety of less traditional outcomes. Given the breadth and depth of policies negotiated in PTAs and their far-reaching effects it is expected and desirable that these agreements continue to be carefully scrutinized both by its citizens and multilateral organizations.

ACKNOWLEDGMENTS

I am very grateful for comments provided by the editors, Kyle Bagwell and Robert Staiger, as well as Stephanie Aaronson, Emily Blanchard, Chad Bown, Kyle Handley, Rob Johnson, Pravin Krishna, Michele Ruta, Emanuel Ornelas, Alan Winters, and participants at the Handbook of Commercial Policy Conference. Marisol Rodriguez-Chatruc and Edith Laget provided excellent research assistance. I acknowledge financial support from the NSF under grant SES-1360780.

REFERENCES

- Abrego, L.E., Perroni, C., Whalley, J., Wigle, R.M., 2001. Trade and environment: bargaining outcomes from linked negotiations. *Rev. Int. Econ.* 9 (3), 414–428.
- Aghion, P., Antràs, P., Helpman, E., 2007. Negotiating free trade. *J. Int. Econ.* 73 (1), 1–30.
- Aguirregabiria, V., Suzuki, J., 2016. Empirical games of market entry and spatial competition in retail industries. In: Basker, E. (Ed.), *Handbook on the Economics of Retailing and Distribution*. Vol. 201. Edward Elgar Publishing, Northampton, MA, pp. 201–232. Chapter 9.
- Amador, M., Bagwell, K., 2013. The theory of optimal delegation with an application to tariff caps. *Econometrica* 81 (4), 1541–1599.
- Anderson, J.E., Van Wincoop, E., 2003. Gravity with gravitas: a solution to the border puzzle. *Am. Econ. Rev.* 93 (1), 170–192.
- Anderson, J.E., Yotov, Y.V., 2016. Terms of trade and global efficiency effects of free trade agreements, 1990–2002. *J. Int. Econ.* 99, 279–298.

- Bagwell, K., Staiger, R.W., 1990. A theory of managed trade. *Am. Econ. Rev.* 80 (4), 779–795.
- Bagwell, K., Staiger, R.W., 1997a. Multilateral tariff cooperation during the formation of customs unions. *J. Int. Econ.* 42 (1), 91–123.
- Bagwell, K., Staiger, R.W., 1997b. Multilateral tariff cooperation during the formation of free trade areas. *Int. Econ. Rev.* 38, 291–319.
- Bagwell, K., Staiger, R.W., 1998. Regionalism and multilateral tariff cooperation. In: Piggott, J., Woodland, A. (Eds.), *International Trade Policy and the Pacific Rim*. MacMillan, London.
- Bagwell, K., Staiger, R.W., 2002. *The Economics of the World Trading System*. MIT Press, Cambridge, MA.
- Bagwell, K., Staiger, R.W., 2011. What do trade negotiators negotiate about? Empirical evidence from the World Trade Organization. *Am. Econ. Rev.* 101 (4), 1238–1273.
- Bagwell, K., Staiger, R.W., 2016. The design of trade agreements. In: Bagwell, K., Staiger, R. (Eds.), *Handbook of Commercial Policy*, vol. 1A. Elsevier, Amsterdam, North-Holland, pp. 435–529.
- Bagwell, K., Bown, C.P., Staiger, R.W., forthcoming. Is the WTO passé? *J. Econ. Liter.*
- Baier, S.L., Bergstrand, J.H., 2004. Economic determinants of free trade agreements. *J. Int. Econ.* 64 (1), 29–63.
- Baier, S.L., Bergstrand, J.H., 2007. Do free trade agreements actually increase members' international trade? *J. Int. Econ.* 71 (1), 72–95.
- Baier, S.L., Bergstrand, J.H., Feng, M., 2014. Economic integration agreements and the margins of international trade. *J. Int. Econ.* 93 (2), 339–350.
- Baldwin, R., 1995. A domino theory of regionalism. In: Baldwin, R.E., Haaparanta, P., Kiander, J. (Eds.), *Expanding Membership in the European Union*. University Press, Cambridge.
- Baldwin, R., 2016. The World Trade Organization and the future of multilateralism. *J. Econ. Perspect.* 30 (1), 95–116.
- Baldwin, R., Jaimovich, D., 2012. Are free trade agreements contagious? *J. Int. Econ.* 88 (1), 1–16.
- Baldwin, R.E., Magee, C.S., 2000. Is trade policy for sale? Congressional voting on recent trade bills. *Public Choice* 105 (1–2), 79–101.
- Baldwin, R., Rieder, R., 2007. A test of endogenous trade bloc formation theory on EU data. CEPR Discussion Papers 6389.
- Baldwin, R.E., Venables, A.J., 1995. Regional economic integration. *Handb. Int. Econ.* 3, 1597–1644.
- Baltagi, B.H., Egger, P., Pfaffermayr, M., 2008. Estimating regional trade agreement effects on FDI in an interdependent world. *J. Econ.* 145 (1), 194–208.
- Beaulieu, E., 2002. Factor or industry cleavages in trade policy? An empirical analysis of the Stolper–Samuelson theorem. *Econ. Polit.* 14 (2), 99–131.
- Bergsten, C.F., 1996. Globalizing free trade. *Foreign Affairs*, 105–120.
- Bergstrand, J.H., 1985. The gravity equation in international trade: some microeconomic foundations and empirical evidence. *Rev. Econ. Stat.* 67 (3), 474–481.
- Beshkar, M., Bond, E.W., 2015. Cap and Escape in Trade Agreements. Indiana University, IN. Unpublished manuscript.
- Beshkar, M., Bond, E.W., Rho, Y., 2015. Tariff binding and overhang: theory and evidence. *J. Int. Econ.* 97, 1–13.
- Bhagwati, J., 1991. *The World Trading System at Risk*. Princeton University Press, Princeton, NJ.
- Bhagwati, J., Krishna, P., Panagariya, A. (Eds.), 1999. *Trading Blocs: Alternative Approaches to Analyzing Preferential Trade Agreements*. MIT Press, Cambridge, MA.
- Blanchard, E.J., 2007. Foreign direct investment, endogenous tariffs, and preferential trade agreements. *BEJ. Econ. Anal. Policy* 7 (1), 1–52.
- Blanchard, E., Matschke, X., 2015. Us multinationals and preferential market access. *Rev. Econ. Stat.* 97 (4), 839–854.
- Blanchard, E.J., Bown, C.P., Johnson, R.C., 2016. Global supply chains and trade policy. NBER Working Paper 21883.
- Blomstrom, M., Kokko, A., 1997. Regional integration and foreign direct investment. NBER Working Paper 6019.
- Bohara, A., Gawande, K., Sanguinetti, P., 2004. Trade diversion and declining tariffs: evidence from Mercosur. *J. Int. Econ.* 64 (1), 65–88.

- Bond, E.W., Syropoulos, C., 1996. The size of trading blocs market power and world welfare effects. *J. Int. Econ.* 40 (3), 411–437.
- Bown, C.P., 2011. Taking stock of antidumping, safeguards and countervailing duties, 1990–2009. *World Econ.* 34 (12), 1955–1998.
- Bown, C., Crowley, M., 2013. Import protection, business cycles, and exchange rates: evidence from the great recession. *J. Int. Econ.* 90 (1), 50–64.
- Bown, C.P., Crowley, M.A., 2016. The empirical landscape of trade policy. In: Bagwell, K., Staiger, R. (Eds.), *Handbook of Commercial Policy*, vol. 1A. Elsevier, Amsterdam, North-Holland, pp. 3–108.
- Bown, C., Karacaovali, B., Tovar, P., 2014. What do we know about preferential trade agreements and temporary trade barriers? In: Dür, A., Elsig, M. (Eds.), *Trade Cooperation: The Purpose, Design and Effects of Preferential Trade Agreements*. Cambridge University Press, Cambridge, UK, pp. 433–462. **Chapter 18.**
- Broda, C., Limão, N., Weinstein, D.E., 2008. Optimal tariffs and market power: the evidence. *Am. Econ. Rev.* 98 (5), 2032–2065.
- Bustos, P., 2011. Trade liberalization, exports, and technology upgrading: evidence on the impact of MERCOSUR on Argentinian firms. *Am. Econ. Rev.* 101 (1), 304–340.
- Cadot, O., Olarreaga, M., Tschopp, J., 2010. Trade agreements and trade barrier volatility. In: Anderson, K. (Ed.), *The Political Economy of Agricultural Price Distortions*. Cambridge University Press, Cambridge, United Kingdom, pp. 332–357. **Cambridge Books Online.**
- Caliendo, L., Parro, F., 2015. Estimates of the trade and welfare effects of NAFTA. *Rev. Econ. Stud.* 82 (1), 1–44.
- Caliendo, L., Feenstra, R.C., Romalis, J., Taylor, A.M., 2015. Tariff reductions, entry, and welfare: theory and evidence for the last two decades. **NBER Working Paper No. 21768.**
- Carballo, J.R., Handley, K., Limão, N., 2015. Trade collapse: the role of economic and policy uncertainty in the Great recession. In: *Essays in Trade and Uncertainty*. PhD dissertation, University of Maryland.
- Chaney, T., 2008. Distorted gravity: the intensive and extensive margins of international trade. *Am. Econ. Rev.* 98 (4), 1707–1721.
- Chang, P.-L., Lee, M.-J., 2011. The WTO trade effect. *J. Int. Econ.* 85 (1), 53–71.
- Chang, W., Winters, A.L., 2002. How regional blocs affect excluded countries: the price effects of Mercosur. *Am. Econ. Rev.* 92 (4), 889–904.
- Chen, M.X., 2009. Regional economic integration and geographic concentration of multinational firms. *Eur. Econ. Rev.* 53 (3), 355–375.
- Chen, M.X., Mattoo, A., 2008. Regionalism in standards: good or bad for trade? *Can. J. Econ.* 41 (3), 838–863.
- Cipollina, M., Salvatici, L., 2010. Reciprocal trade agreements in gravity models: a meta-analysis. *Rev. Int. Econ.* 18 (1), 63–80.
- Clausing, K.A., 2001. Trade creation and trade diversion in the Canada–United States free trade agreement. *Can. J. Econ.* 34 (3), 677–696.
- Conconi, P., Facchini, G., Zanardi, M., 2014. Policymakers’ horizon and trade reforms: the protectionist effect of elections. *J. Int. Econ.* 94 (1), 102–118.
- Costinot, A., Rodriguez-Clare, A., 2014. Trade theory with numbers: quantifying the consequences of globalization. *Handb. Int. Econ.* 4, 197.
- Crivelli, P., 2014. Regionalism and falling external protection in high and low tariff members. **Working Paper 14–08–2**, Geneva School of Economics and Management.
- Deason, L., 2014. Modelling international trade under the threat of tariff hikes in general equilibrium. PhD dissertation, University of Maryland, College Park, MD.
- Dür, A., Baccini, L., Elsig, M., 2014. The design of international trade agreements: introducing a new dataset. *Rev. Int. Organ.* 9 (3), 353–375.
- Dutt, P., Mihov, I., Van Zandt, T., 2013. The effect of WTO on the extensive and the intensive margins of trade. *J. Int. Econ.* 91 (2), 204–219.
- Ederington, J., Ruta, M., 2016. Nontariff measures and the world trading system. In: Bagwell, K., Staiger, R. (Eds.), *Handbook of Commercial Policy*, vol 1B. Elsevier, Amsterdam, North-Holland, pp. 211–277.

- Egger, P., Larch, M., 2008. Interdependent preferential trade agreement memberships: an empirical analysis. *J. Int. Econ.* 76 (2), 384–399.
- Egger, P., Wamser, G., 2013. Effects of the endogenous scope of preferentialism on international goods trade. *BE J. Econ. Anal. Policy*, 13 (2), 709–731.
- Egger, H., Egger, P., Greenaway, D., 2008. The trade structure effects of endogenous regional trade agreements. *J. Int. Econ.* 74 (2), 278–298.
- Egger, P., Larch, M., Staub, K.E., Winkelmann, R., 2011. The trade effects of endogenous preferential trade agreements. *Am. Econ. J. Econ. Policy*, 3 (3), 113–143.
- Egger, P., Wamser, G., Beck, T., Conconi, P., 2013. Multiple faces of preferential market access: their causes and consequences. *Econ. Policy* 28 (73), 143–187.
- Eichengreen, B., O’rourke, K.H., 2009. A tale of two depressions, Available at <http://voxeu.org/article/tale-two-depressions-what-do-new-data-tell-us-february-2010-update>.
- Eicher, T.S., Henn, C., 2011. In search of WTO trade effects: preferential trade agreements promote trade strongly, but unevenly. *J. Int. Econ.* 83 (2), 137–153.
- Eicher, T.S., Henn, C., Papageorgiou, C., 2012. Trade creation and diversion revisited: accounting for model uncertainty and natural trading partner effects. *J. Appl. Econ.* 27 (2), 296–321.
- Estevadeordal, A., Freund, C., Ornelas, E., 2008. Does regionalism affect trade liberalization toward non-members? *Q. J. Econ.* 123 (4), 1531–1575.
- Feenstra, R.C., 1989. Symmetric pass-through of tariffs and exchange rates under imperfect competition: an empirical test. *J. Int. Econ.* 27 (1), 25–45.
- Fernandez, R., Portes, J., 1998. Returns to regionalism: an analysis of nontraditional gains from regional trade agreements. *World Bank Econ. Rev.* 12 (2), 197–220.
- Foroutan, F., 1998. Does membership in a regional preferential trade arrangement make a country more or less protectionist? *World Econ.* 21 (3), 305–336.
- Francois, J., Manchin, M., Norberg, H., Pindyuk, O., Tomberger, P., 2013. Reducing transatlantic barriers to trade and investment: an economic assessment. IIDE Discussion Papers 20130401, Institute for International and Development Economics.
- Frankel, J.A., Stein, E., Wei, S.-J., 1995. Trading blocs and the Americas: the natural, the unnatural, and the super-natural. *J. Dev. Econ.* 47 (1), 61–95.
- Frankel, J.A., Stein, E., Wei, S.-J., 1997. Regional Trading Blocs in the World Economic System. Peterson Institute for International Economics, Washington, DC.
- Freund, C.L., Ornelas, E., 2010. Regional trade agreements. *Annu. Rev. Econ.* 2, 139–166.
- Fugazza, M., Robert-Nicoud, F., 2014. The ‘emulator effect’ of the Uruguay round on US regionalism. *Rev. Int. Econ.* 22 (5), 1049–1078.
- Furusawa, T., Konishi, H., 2007. Free trade networks. *J. Int. Econ.* 72 (2), 310–335.
- Ghosh, S., Yamarik, S., 2004. Are regional trading arrangements trade creating? *J. Int. Econ.* 63 (2), 369–395.
- Goldberg, P.K., Pavcnik, N., 2016. The effects of trade policy. In: Bagwell, K., Staiger, R. (Eds.), *Handbook of Commercial Policy*, vol. 1A. Elsevier, Amsterdam, North-Holland, pp. 161–206.
- Goyal, S., Joshi, S., 2006. Bilateralism and free trade. *Int. Econ. Rev.* 47 (3), 749–778.
- Grosso, V., Piermartini, R., 2014. Trade policy uncertainty and the WTO. DIW Berlin Discussion Paper 1437.
- Grossman, G.M., 2016. The purpose of trade agreements. In: Bagwell, K., Staiger, R. (Eds.), *Handbook of Commercial Policy*, vol. 1A. Elsevier, Amsterdam, North-Holland, pp. 379–434.
- Grossman, G.M., Helpman, E., 1994. Protection for sale. *Am. Econ. Rev.* 84 (4), 833–850.
- Grossman, G., Helpman, E., 1995. The politics of free trade agreements. *Am. Econ. Rev.* 85 (4), 667–690.
- Hafner-Burton, E.M., 2013. *Forced to Be Good: Why Trade Agreements Boost Human Rights*. Cornell University Press, Ithaca, NY.
- Hakobyan, S., McLaren, J., 2010. Looking for local labor market effects of NAFTA. NBER Working Paper 16535.
- Handley, K., 2014. Exporting under trade policy uncertainty: theory and evidence. *J. Int. Econ.* 94 (1), 50–66.
- Handley, K., Limão, N., 2013. Policy uncertainty, trade and welfare: theory and evidence for China and the US. NBER Working Paper 19376.

- Handley, K., Limão, N., 2015. Trade and investment under policy uncertainty: theory and firm evidence. *Am. Econ. J. Policy* 7 (4), 189–222.
- Hayakawa, K., Kimura, F., 2014. How much do free trade agreements reduce impediments to trade? *Open Econ. Rev.* 26 (4), 1–19.
- Head, K., Mayer, T., 2014. Gravity equations: workhorse, toolkit, and cookbook. *Handb. Int. Econ.* 4.
- Head, K., Ries, J., 1999. Rationalization effects of tariff reductions. *J. Int. Econ.* 47 (2), 295–320.
- Hertel, T., Hummels, D., Ivanic, M., Keeney, R., 2007. How confident can we be of CGE-based assessments of free trade agreements? *Econ. Model.* 24 (4), 611–635.
- Hoekman, B., Kostecki, M., 2001. *The Political Economy of the World Trading System*. Oxford University Press, Oxford, United Kingdom.
- Horn, H., Mavroidis, P.C., Sapir, A., 2010a. Beyond the WTO? An anatomy of EU and US preferential trade agreements. *World Econ.* 33 (11), 1565–1588.
- Horn, H., Maggi, G., Staiger, R.W., 2010b. Trade agreements as endogenously incomplete contracts. *Am. Econ. Rev.* 100 (1), 394–419.
- Imbens, G.W., 2004. Nonparametric estimation of average treatment effects under exogeneity: a review. *Rev. Econ. Stat.* 86 (1), 4–29.
- Jackson, J.H., 1997. *The World Trading System: Law and Policy of International Economic Relations*, second ed. MIT Press, Cambridge, MA. xi, p. 441.
- Johnson, H.G., 1967. *International Trade and Economic Growth*. George Allen and Unwin, London, United Kingdom.
- Johnson, R.C., Noguera, G., 2014. A portrait of trade in value added over four decades. Unpublished Manuscript, Dartmouth College.
- Karacaovali, B., Limão, N., 2008. The clash of liberalizations: preferential vs. multilateral trade liberalization in the European Union. *J. Int. Econ.* 74 (2), 299–327.
- Kee, H.L., Olarreaga, M., Silva, P., 2007. Market access for sale. *J. Dev. Econ.* 82 (1), 79–94.
- Kee, H.L., Nicita, A., Olarreaga, M., 2009. Estimating trade restrictiveness indices. *Econ. J.* 119 (534), 172–199.
- Kee, H.L., Neagu, C., Nicita, A., 2013. Is protectionism on the rise? Assessing national trade policies during the crisis of 2008. *Rev. Econ. Stat.* 95 (1), 342–346.
- Kehoe, T.J., 2005. An evaluation of the performance of applied general equilibrium models of the impact of NAFTA. In: Kehoe, T.J., Srinivasan, T.N., Whalley, J. (Eds.), *Frontiers in Applied General Equilibrium Modeling: Essays in Honor of Herbert Scarf*. Cambridge University Press, Cambridge, UK.
- Kemp, M.C., Wan, H.Y., 1976. An elementary proposition concerning the formation of customs unions. *J. Int. Econ.* 6 (1), 95–97.
- Ketterer, T., Bernhofen, D., Milner, C., 2014. Preferences, rent destruction and multilateral liberalization: the building block effect of CUSFTA. *J. Int. Econ.* 92 (1), 63–77.
- Ketterer, T., Bernhofen, D., Milner, C., 2015. The impact of trade preferences on multilateral tariff cuts: evidence for Japan. *J. Jpn. Int. Econ.* 38, 31–51.
- Kreinin, M.E., 1961. Effect of tariff changes on the prices and volume of imports. *Am. Econ. Rev.* 51 (3), 310–324.
- Krishna, P., 1998. Regionalism and multilateralism: a political economy approach. *Q. J. Econ.* 113 (1), 227–251.
- Krishna, P., 2003. Are regional trading partners “natural”? *J. Polit. Econ.* 111 (1), 202–226.
- Krishna, P., 2008. The economics of preferential trade agreements. In: Kwan Choi, E., Hartigan, J.C. (Eds.), *Handbook of International Trade: Economic and Legal Analyses of Trade Policy and Institutions*. Vol. II. Blackwell Publishing Ltd, Hoboken, NJ, pp. 294–312. Chapter 9.
- Krugman, P., 1991. The move to free trade zones. In: *Policy Implications of Trade and Currency Zones: A Symposium*. Fed. Reserve Bank, Kansas City, MO.
- Levchenko, A.A., Zhang, J., 2012. Comparative advantage and the welfare impact of European integration. *Econ. Policy* 27 (72), 567–602.
- Levy, P.I., 1997. A political-economic analysis of free-trade agreements. *Am. Econ. Rev.* 87 (4), 506–519.

- Levy-Yeyati, E.L., Stein, E., Daude, C., 2003. Regional integration and the location of FDI. IDB Working Paper 492.
- Lileeva, A., Trefler, D., 2010. Improved access to foreign markets raises plant-level productivity... for some plants. *Q. J. Econ.* 125 (3), 1051–1099.
- Limão, N., 2005. Trade policy, cross-border externalities and lobbies: do linked agreements enforce more cooperative outcomes? *J. Int. Econ.* 67 (1), 175–199.
- Limão, N., 2006. Preferential trade agreements as stumbling blocks for multilateral trade liberalization: evidence for the United States. *Am. Econ. Rev.* 96 (3), 896–914.
- Limão, N., 2007. Are preferential trade agreements with non-trade objectives a stumbling block for multilateral liberalization? *Rev. Econ. Stud.* 74 (3), 821–855.
- Limão, N., Maggi, G., 2015. Uncertainty and trade agreements. *Am. Econ. J. Microecon* 7 (4), 1–42.
- Limão, N., Olarreaga, M., 2006. Trade preferences to small developing countries and the welfare costs of lost multilateral liberalization. *World Bank Econ. Rev.* 20 (2), 217–240.
- Limão, N., Tovar, P., 2011. Policy choice: theory and evidence from commitment via international trade agreements. *J. Int. Econ.* 85 (2), 186–205.
- Lipsey, R.G., Lancaster, K., 1957. The general theory of second best. *Rev. Econ. Stud.* 24 (1), 11–32.
- Liu, X., Ornelas, E., 2014. Free trade agreements and the consolidation of democracy. *Am. Econ. J. Macroecon.* 6 (2), 29–70.
- Ludema, R.D., Mayda, A.M., 2009. Do countries free ride on MFN? *J. Int. Econ.* 77 (2), 137–150.
- Ludema, R.D., Mayda, A.M., 2013. Determs-of-trade effects matter for trade agreements? Evidence from WTO countries. *Q. J. Econ.* 128 (4), 1837–1893.
- Machlup, F., 1977. Publishing scholarly books and journals: is it economically viable? *J. Polit. Econ.* 85 (1), 217–225.
- Madsen, J.B., 2001. Trade barriers and the collapse of world trade during the great depression. *South. Econ. J.* 67 (4), 848–868.
- Magee, C.S., 2003. Endogenous preferential trade agreements: an empirical analysis. *Contrib. Econ. Anal. Policy* 2 (1), 1–19.
- Magee, C.S., 2015. Trade creation, trade diversion, and the general equilibrium effects of regional trade agreements: a study of the European Community–Turkey customs union. *Rev. World Econ.* 152 (2), 383–399.
- Maggi, G., 1999. The role of multilateral institutions in international trade cooperation. *Am. Econ. Rev.* 89 (1), 190–214.
- Maggi, G., 2014. International trade agreements. *Handb. Int. Econ.* 4, 317.
- Maggi, G., 2016. Issue linkage. In: Bagwell, K., Staiger, R. (Eds.), *Handbook of Commercial Policy*, vol. 1B. Elsevier, Amsterdam, North-Holland, pp. 513–564.
- Maggi, G., Rodriguez-Clare, A., 1998. The value of trade agreements in the presence of political pressures. *J. Polit. Econ.* 106 (3), 574–601.
- Mansfield, E.D., Milner, H.V., 2012. *Votes, Vetoes, and the Political Economy of International Trade Agreements*. Princeton University Press, Princeton, NJ.
- Martin, P., Mayer, T., Thoenig, M., 2008. Make trade not war? *Rev. Econ. Stud.* 75 (3), 865–900.
- Martin, P., Mayer, T., Thoenig, M., 2012. The geography of conflicts and regional trade agreements. *Am. Econ. J. Macroecon.* 4 (4), 1–35.
- McLaren, J., 2016. The political economy of commercial policy. In: Bagwell, K., Staiger, R. (Eds.), *Handbook of Commercial Policy*, vol 1A. Elsevier, Amsterdam, North-Holland, pp. 110–159.
- Mitra, D., Thomakos, D.D., Ulubasoglu, M.A., 2002. ‘Protection for sale’ in a developing country: democracy vs. dictatorship. *Rev. Econ. Stat.* 84 (3), 497–508.
- Novy, D., 2013. Gravity redux: measuring international trade costs with panel data. *Econ. Inq.* 51 (1), 101–121.
- Olarreaga, M., Soloaga, I., Alan Winters, L., 1999. What’s Behind MERCOSUR’s Common External Tariff? vol. 2310. World Bank Publications, Washington, DC.
- Orefice, G., Rocha, N., 2014. Deep integration and production networks: an empirical analysis. *World Econ.* 37 (1), 106–136.

- Ornelas, E., 2005. Rent destruction and the political viability of free trade agreements. *Q. J. Econ.* 120 (4), 1475–1506.
- Ornelas, E., 2016. Special and differential treatment for developing countries. In: Bagwell, K., Staiger, R. (Eds.), *Handbook of Commercial Policy*, vol. 1B. Elsevier, Amsterdam, North-Holland, pp. 369–432.
- Osnago, A., Rocha, N., Ruta, M., 2015. Deep trade agreements and vertical FDI: the devil is in the details. *World Bank Policy Research Working Paper 7464*.
- Ossa, R., 2014. Trade wars and trade talks with data. *Am. Econ. Rev.* 104 (12), 4104–4146.
- Özden, Ç., Reinhardt, E., 2005. The perversity of preferences: GSP and developing country trade policies, 1976–2000. *J. Dev. Econ.* 78 (1), 1–21.
- Panagariya, A., Krishna, P., 2002. On necessarily welfare enhancing free trade areas. *J. Int. Econ.* 57 (2), 353–367.
- Pardo, C., Hector, F., Freund, C.L., Ornelas, E., 2009. The ASEAN free trade agreement: impact on trade flows and external trade barriers. *World Bank Policy Research Working Paper 4960*.
- Prusa, T.J., Teh, R., 2010. Protection reduction and diversion: PTAS and the incidence of antidumping disputes. *NBER Working Paper 16276*.
- Richardson, M., 1993. Endogenous protection and trade diversion. *J. Int. Econ.* 34 (3–4), 309–324.
- Riezman, R., 1985. Customs unions and the core. *J. Int. Econ.* 19 (3), 355–365.
- Romalis, J., 2007. NAFTA's and CUSFTA's impact on international trade. *Rev. Econ. Stat.* 89 (3), 416–435.
- Rose, A.K., 2004. Do we really know that the WTO increases trade? *Am. Econ. Rev.* 94 (1), 98–114.
- Ruhl, K.J., 2004. The international elasticity puzzle. *Working Paper*, NYU Stern School of Business.
- Saggi, K., 2016. Trade, intellectual property rights, and the world trade organization. In: Bagwell, K., Staiger, R. (Eds.), *Handbook of Commercial Policy*, vol. 1B. Elsevier, Amsterdam, North-Holland, pp. 433–512.
- Saggi, K., Yildiz, H.M., 2010. Bilateralism, multilateralism, and the quest for global free trade. *J. Int. Econ.* 81 (1), 26–37.
- Saggi, K., Woodland, A., Yildiz, H.M., 2013. On the relationship between preferential and multilateral trade liberalization: the case of customs unions. *Am. Econ. J. Microecon.* 5 (1), 63–99.
- Sapir, A., 1997. Domino effects in Western European Trade, 1960–1992. *Eur. J. Polit. Econ.* 17, 377–386.
- Spearot, A., 2016. Unpacking the long run effects of tariff shocks: new structural implications from firm heterogeneity models. *AEJ Microecon.* 8 (2), 128–167.
- Staiger, R.W., Tabellini, G., 1987. Discretionary trade policy and excessive protection. *Am. Econ. Rev.* 77 (5), 823–837.
- Subramanian, A., Wei, S.-J., 2007. The WTO promotes trade, strongly but unevenly. *J. Int. Econ.* 72 (1), 151–175.
- Tinbergen, J., 1962. Shaping the world economy: suggestions for an international economic policy. No. HD82 T54.
- Tintelnot, F., 2015. Global production with export platforms. <http://felix-tintelnot.wikidot.com/research>>.
- Tovar, P., 2012. Preferential trade agreements and unilateral liberalization: evidence from CAFTA. *World Trade Rev.* 11 (4), 591–619.
- Trefler, D., 2004. The long and short of the Canada-US free trade agreement. *Am. Econ. Rev.* 94 (4), 870–895.
- United States International Trade Commission (USITC), 2008. *Year in Review*. USITC, Washington, DC.
- Vicard, V., 2009. On trade creation and regional trade agreements: does depth matter? *Rev. World Econ.* 145 (2), 167–187.
- Vicard, V., 2012. Trade, conflicts and political integration: explaining the heterogeneity of regional trade agreements. *Eur. Econ. Rev.* 56 (1), 54–71.
- Viner, J., 1950. The economics of customs unions. In: Oslington, P. (Ed.), *The Customs Union Issue*. Oxford University Press, Oxford, United Kingdom, pp. 41–81.
- Winters, A.L., 1997. Regionalism and the rest of the world: the irrelevance of the Kemp-Wan theorem. *Oxf. Econ. Pap.* 49 (2), 228–234.
- Winters, A.L., 1999. Regionalism vs. multilateralism. In: Baldwin, R., Cohen, D., Sapir, A., Venables, T. (Eds.), *Market Integration, Regionalism and the Global Economy*. CEPR, Cambridge, UK. **Chapter 2**.

- Winters, A.L., Chang, W., 2000. Regional integration and import prices: an empirical investigation. *J. Int. Econ.* 51 (2), 363–377.
- World Bank, 2000. *Regional Integration Agreements*. World Bank, Washington, DC.
- World Bank, 2005. *Global Economic Prospects 2005: Trade, Regionalism, and Development*. World Bank, Washington, DC.
- WTO, World Trade Organisation, 2011. *World trade report 2011—The WTO and preferential trade agreements: from co-existence to coherence*.
- Yi, S.-S., 1996. Endogenous formation of customs unions under imperfect competition: open regionalism is good. *J. Int. Econ.* 41 (1), 153–177.
- Yi, K.-M., 2003. Can vertical specialization explain the growth of world trade? *J. Polit. Econ.* 111 (1), 52–102.