



**GLOBAL FORUM ON TRADE
RECONCILING REGIONALISM AND MULTILATERALISM IN A POST-BALI WORLD**

MULTILATERALISING 21ST CENTURY REGIONALISM

Richard Baldwin
Professor of International Economics,
Graduate Institute Geneva

**OECD CONFERENCE CENTRE, PARIS
11-12 (a.m.) February, 2014**



MULTILATERALISING 21ST CENTURY REGIONALISM

The multilateralisation of regionalism takes different forms when applied to deep versus shallow regional trade agreements (RTAs). Shallow agreements focus on discriminatory tariffs; hence, multilateralisation strives mainly to reduce discrimination. Deep agreements focus on the disciplines necessary to foster international production sharing; key provisions often resembling unilateral liberalisations that just happen to be bound by an RTA. In this case, multilateralisation achieves network externalities and solves co-ordination problems. This paper suggests a novel framework for thinking about the costs and benefits of multilateralising the provisions in deep RTAs, including those that seem set to appear in the Trans-Pacific Partnership (TPP) and the Trans-Atlantic Trade and Investment Partnership (TTIP).

TABLE OF CONTENTS

1. Introduction	5
2. Defining 20 th and 21 st century regionalism	6
Twentieth-century RTAs	6
Twenty-first century RTAs	9
Not just RTAs: 21st century regionalism	10
3. Economics of 20th century and 21st century trade	17
20th and 21st century comparative advantage	17
Twentieth-century trade: Basic economics	18
Twenty-first century trade: Basic economics	20
4. The impact of regionalism: Empirics	24
Shallow RTAs: Trade creation and diversion	24
Deep RTAs and supply chain trade	25
Summary	27
5. Liberalising 20 th versus 21 st century trade	27
Liberalising 20th century trade: Juggernauts and dominos	27
International political economy of 21st century trade liberalisation	29
6. Multilateralising deep regionalism: Practical issues	30
Multilateralising preferential agreements on services trade	30
Competition policy	31
Investment: RTAs, BITs and capital movement provisions	31
Technical barriers to trade (TBTs)	32
Summary	33
7. Which measures should be multilateralised?	33
“Fiscal federalism” as an analytic framework	33
The lack of spillovers	35
Network externalities and “standards competition” analysis	35
A tentative list of deep RTA provisions to be multilateralised	37
Multi-tier multilateralisation	38
8. Conclusions	39
The real threat to multilateralism	39
Multilateralising 21 st century regionalism: A better way forward	40
The way forward	40
References	41

Tables

Table 1. Margins of preference in 2008	8
Table 2. Example of deep RTA provisions in the WTO database	14

Figures

Figure 1. Share of imports with MFN zero tariffs, various RTAs, 1995 to 2008	7
Figure 2. Disciplines underpinning international production sharing	10
Figure 3. Unilateral tariff liberalisation, 1988 – 2008	11

Figure 4. Take-off in BITs and FDI.....	12
Figure 5. Frequency of legally enforceable	13
Figure 6. Deep RTA provisions and number of RTAs	14
Figure 7. Share of US and Japanese agreements with deeper provisions	15
Figure 8. Share of European Union and rest of world (RoW) agreements with deeper provisions.....	16
Figure 11. Impact of 21st century RTA: Switching comparative advantage	21
Figure 12. Supply chain trade with intermediate goods and no technology lending	23
Figure 13. Recent estimates of trade creation and trade diversion	26
Figure 14. Juggernaut and domino effects	28
Figure 15. Schema for thinking about multilateralisation levels	36

1. Introduction¹

When I coined the phrase “multilateralising regionalism” in the 2006 World Economy Annual Lecture, conventional wisdom viewed preferential trade agreements as focusing mostly on tariffs.² In this setting, multilateralising regionalism could be defined as making regional trade agreements (RTAs) less preferential. Multilateralisation meant turning systems of bilateral RTAs into regional zones of duty-free trade as a step to removing tariffs globally. The focus was on tariffs, rules of origin and rules of cumulation (Baldwin and Low, 2009). Multilateralising 21st century regionalism is a very different thing.

Many 21st century RTAs include provisions unrelated to tariffs or other border measures. This is because 21st century RTAs are at the same time trade agreements and production-sharing agreements. The trade aspects reduce barriers to selling foreign-made goods. The production-sharing aspects lock in disciplines that facilitate the internationalisation of production – especially between high-tech and low-wage nations. At first blush, this may sound like 21st century regionalism is “offshoring jobs”, but recent empirical work shows this can be a win-win situation (Hufbauer, Moran, and Oldenski, 2013). Expanded activity of affiliates abroad is associated with greater production, employment and research and development (R&D) in the home nation.

Goals of this paper

The goals of this paper are twofold. The first is to argue that multilateralising 21st century regionalism would be a good way to ensure that the ongoing mega-regional talks are constructive steps toward improving the multilateral trade governance system. This point is taken as largely self-evident and receives little room in the paper.

The second goal is to argue that we need a different mindset when thinking about 21st century multilateralisation. This point requires a good deal of background, as shifting a mindset always does.

Our first step is to argue that 20th and 21st century regionalism are fundamentally different. Twentieth-century RTAs concern “made-here-sold-there” goods, while 21st century regionalism concerns “made-everywhere-sold-there” goods. The difference means that 21st century RTAs include rules on *making* goods as well as *selling* them. These rules impinge upon firms, services, capital, regulations and intellectual property (IP). The paper argues that discrimination is technically difficult for such rules, since it is hard to define the nationality of firms, services, capital, and IP in ways that cannot be easily circumvented. For this reason and others, 21st century regionalism is not fundamentally about discrimination. It is about undergirding the internationalisation of production processes.

Our second step is to argue that the economic effects of global value chain (GVC)-linked trade, i.e. 21st century trade, differ fundamentally from those of made-here-sold-there trade, i.e. 20th century trade. GVC-linked trade changes nations’ comparative advantage, since it de-nationalises the whole notion of comparative advantage. The competitiveness of GVC-produced goods depends upon a multinational bundle of labour, capital and technology. Under the 20th century conceptualisation of trade, by contrast, production is national, so trade involves competition among national bundles of capital, labour and technology. This difference explains why GVC participation is now the fast-track to industrial development. Joining a GVC allows nations to export goods they never could on their own.

Our next step is to argue that the distinctions between 20th and 21st century trade and regionalism require a new mindset. The old paradigms are inadequate for thinking through the new challenges. Twentieth-century concepts like “trade creation and diversion”, “spaghetti bowls” and “building and stumbling blocks” were relevant when regionalism was mostly about discrimination. Now that regionalism is largely about underpinning international production networks, the old concepts are unhelpful in most cases and harmful in others.

-
1. This paper was written for the Organisation for Economic Trade and Co-operation (OECD) in the context of its assessment of whether and how the World Trade Organization (WTO)-plus steps taken in RTAs could be harnessed to multilateralise liberalisation and rule-making. My thanks to Susan Stone for her valuable comments and to Yuan Zi for excellent research assistance
 2. Published as Baldwin (2006a) and NBER Working Paper 12545.

Our final step is to offer a fresh approach to thinking about multilateralising 21st century regionalism. I suggest we think of the benefits from 21st century multilateralisation as stemming from the extra “network externalities” that come from knitting together diverse sets of disciplines. Here, “network externalities” means that each individual’s gain from participating in a network increases with the network’s size. I suggest that we think of the costs of multilateralisation in terms of the cost of harmonisation, i.e. in terms of systems competition. For example, if all bilateral investment treaties (BITs) were harmonised, and all property rights assured by the same rules, network externalities would be maximised. But which set of investment rules would be adopted? US firms would surely prefer the US template, while European firms would prefer the European template. This suggests that the cost of harmonisation can be thought of as a “systems competition”. Here, “systems competition” means the sort of problems that would arise if one had to choose whether the Windows, OS or Linux operating system should become the global standard.

As part of this fresh perspective, a new question – the level of governance – arises. As 20th century multilateralisation aimed to limit tariff discrimination, the logic of non-discrimination meant that the multilateral level was the best place to do so. Twentieth-century regionalism, however, involves harmonisations on a far broader range of policies. The key questions are: which deep RTA provisions should be harmonised at the global level and which at the regional level? Which disciplines are best left un-harmonised?

Answering these questions will require a major legal, economic and political research effort to learn more about the exact differences among existing deep RTA provisions and the difficulty of partially or fully harmonising them. As an analogy, the Organisation for Economic Co-operation and Development (OECD) spent decades documenting – in a harmonised manner – the distortionary impact of nations’ radically different agricultural policies. These results laid the groundwork for the Uruguay Round’s successful negotiations on the subject. Without this research, diplomats would have argued incessantly over the basic terms and effects of different nation’s agricultural policies. The OECD played the fair broker in that case. The WTO could do the same in the case of deep RTA provisions as a means of smoothing the road to multilateralisation in the medium term.

2. Defining 20th and 21st century regionalism

This section defines and illustrates the development of 20th and 21st century regionalism. We begin by defining 20th and 21st century trade, a topic to which we return in depth in Section 0.

In a nutshell, 20th century trade is about “made-here-sold-there” goods. In this world, international commerce means goods crossing borders. Twenty-first century trade is about “made-everywhere-sold-there” goods. International commerce thus involves 20th century trade, *plus* complex cross-border flows related to international production networks. It includes trade in intermediate goods, services, ideas, know-how, capital and people.

Twentieth-century RTAs

As twentieth-century trade was mostly about goods crossing borders, twentieth-century RTAs were mainly about trade barriers at the border - especially tariff preferences and related rules (of origin, cumulation, etc.). From the time free trade agreements (FTAs) were defined in Article 24 of the 1947 General Agreement on Tariffs and Trade (GATT), their form has changed little. An FTA sets tariffs to zero on substantially all trade between the signatories. A customs union goes further, harmonising tariffs against third nations. Apart from some rare examples, like the European Economic Community (EEC), pre-1980s RTAs were of this 20th century type. To avoid awkward prose, we refer to all forms of non-multilateral agreements as either preferential trade agreements (PTAs) or RTAs.³

3. Some authors prefer the more logically inclusive PTA, but well-informed observers know that in WTO jargon this refers only to FTAs among developing nations. Moreover, as shown below, tariff preferences have eroded to the point where having “preference” as the key noun is misleading too.

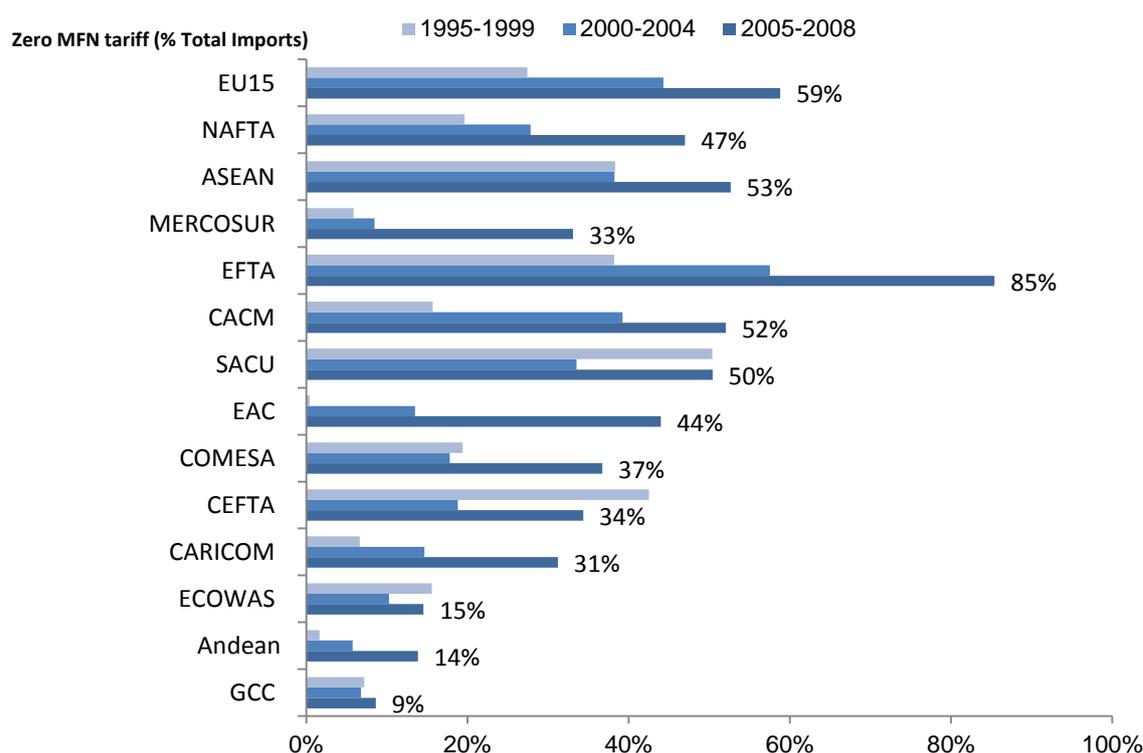
The marginalisation of margins of preferences

Since 20th century regionalism is characterised by tariff preferences, our first step is to study the manner in which preferences have evolved. This section documents how many so-called PTAs are no longer very preferential.

The margin of preference created by an RTA is the difference between the tariff applied to imports from RTA partners as opposed to non-RTA partners, i.e. the countries' most favoured nation (MFN) tariffs. Tariff reductions in advanced economies (driven by GATT Rounds) lowered MFN tariffs to quite low levels, with the result that margins of preference automatically fell. As we shall see below, developing countries' MFN tariffs have dropped more recently, although largely outside of GATT Rounds. This means that the scope for 20th century RTAs to create large tariff preferences is now greatly eroded.

More precisely, Acharya et al. (2011) show that the share of RTA imports that enjoy MFN zero tariffs has risen steadily (Figure 1). Since such products cannot include tariff preference, its importance compared to the early post-war period is greatly diminished. Another excellent study illustrating these basic facts, Fugazza and Nicita (2010), goes one step further by considering interactions between overlapping preference margins. The authors' basic insight is that no one has preferences when everyone has preferences. They also consider import elasticities to determine whether the large preference margins fall on goods whose quantity reacts strongly to small price differences. Despite this refinement, they reach the same conclusion, i.e. that tariff preferences are now rather small from a global perspective.

Figure 1. Share of imports with MFN zero tariffs, various RTAs, 1995 to 2008



Source: Acharya et al. (2011).

Table 1. Margins of preference in 2008

	Share of imports according to margin of preference					Imports (trillion)
	Over 20%	20% to 10%	10% to 5%	Positive but under 5%	Zero preference	
World	1%	2%	7%	18%	69%	USD 13.6
World (ex-intra-European Union)	1%	1%	4%	11%	83%	USD 9.8
Largest importers (over USD 500 billion)						
European Union (internal)	4%	5%	17%	38%	34%	USD 3.8
European Union (external)	0%	2%	3%	11%	82%	USD 2.3
United States	1%	1%	2%	22%	74%	USD 2.1
China	0%	0%	2%	4%	93%	USD 1.0
Japan	0%	0%	1%	5%	93%	USD 0.7
Other top traders						
Mexico	6%	10%	31%	1%	48%	USD 0.30
Canada	0%	2%	26%	8%	65%	USD 0.37
Chile	1%	3%	9%	40%	46%	USD 0.18
Turkey	0%	2%	11%	27%	59%	USD 0.19
Brazil	3%	4%	4%	1%	88%	USD 0.17
Russia	1%	3%	2%	8%	85%	USD 0.19
Indonesia	1%	1%	3%	20%	73%	USD 0.07
Malaysia	1%	2%	1%	1%	92%	USD 0.14
Thailand	1%	1%	1%	4%	93%	USD 0.13
Australia	0%	0%	1%	12%	86%	USD 0.19
Korea	0%	0%	1%	8%	90%	USD 0.43
India	0%	0%	1%	4%	93%	USD 0.22
Singapore	0%	0%	0%	0%	100%	USD 0.24
Taipei, China	0%	0%	0%	0%	100%	USD 0.23
Argentina	0%	0%	0%	5%	95%	USD 0.15
Hong Kong	0%	0%	0%	0%	100%	USD 0.37

Source: Author's calculations, based on Carpenter and Lendle (2010) data.

Carpenter and Lendle (2010) provide even more direct evidence for the 20 largest trading nations. They study tariff line data carefully for actual preferences granted. This is important, since i) many of the tariff lines have applied MFN rates of zero and hence no preference is possible and ii) where MFN tariffs are high, the goods are often excluded from RTAs. As a result, the degree of preferences is radically lower than aggregate numbers might suggest.

Around half of world imports are covered by an RTA. However, only 16.7% of world trade is eligible for preferences⁴. Moreover, the preference margins are low: less than 2% of world imports enjoy preferences over 10 percentage points. These numbers do not consider trade inside the largest RTA of all, the European Union. Taking world totals to include intra-European Union flows, Carpenter and Lendle (2010) calculate that 64% of world trade is covered by an RTA and 29.8% of world trade is subject to preference margins, with only 3.9% enjoying margins over 10 percentage points.

As Table 1 shows, the largest importers' imports are not subject to large preference margins. Intra-European Union trade is by far the most preferential, with 9% carrying preference margins over 10 percentage points. The United States grants preferences over 10 percentage points on 2% of its imports, and China and Japan grant such preference margins to none of their imports. Small nations that are heavily dependent on large neighbours register the highest share of imports covered by margins over 10%. For instance, 16% of Mexican and 2% of Canadian imports receive such margins.

Twenty-first century RTAs

Twenty-first century RTAs, or “deep” RTAs, are quite different. As mentioned above, all of them include tariff preferences, but they are not primarily about preferential market access. Rather, they focus on disciplines underpinning international supply chains. It is useful to distinguish two aspects of international production sharing, each of which creates a need for new types of disciplines (Figure 2). The first is related to:

- Co-ordinating internationally dispersed production facilities.

The disciplines necessary to assure this can be thought of as “supply chain disciplines.” The point is that bringing high-quality, competitively priced goods to customers in a timely manner requires international co-ordination of production facilities via the continuous two-way flow of goods, people, ideas and investments.⁵ Certain policies or national practices threaten these flows, so 21st century RTAs include provisions to restrict such policies.

The second is related to:

- Producing abroad.

The disciplines that underpin this can be thought of as “offshoring disciplines”. When firms set up production facilities abroad, or form long-term ties with foreign suppliers, they typically expose their capital and technical, managerial and marketing know-how to new international risks.⁶ Policies that reduce or eliminate risk to these forms of tangible and intangible property are typically included in 21st century RTAs.

It is vital to remember that these disciplines are a package. All of them are necessary for offshoring firms to feel confident combining their technology with labour in the offshore destination – typically a developing

-
4. The remaining trade flows either have zero MFN tariffs (about 25% of world trade) so there can be no preference, or are excluded from preferential treatment by the terms of the RTA (about 9% of world trade). Applied MFN tariffs are zero for 56% of European Union external trade, 43% for the United States, 48% for China and 80% for Japan. Products for which the large importers maintain high tariffs – especially agricultural goods for developed nations – are routinely excluded from their RTAs.
 5. Tariffs on imported intermediates are one part of this. Co-ordinating international production also requires assurances of world-class telecommunications, goods transportation (especially express parcel services and air cargo) and customs clearance; assured access for short-term visits by key personnel (managers and technicians); and capital and financial market openness to inward and outward investment flows and profit repatriation.
 6. As the World Bank (2011) notes, doing business abroad implicates “the laws, regulations and institutional arrangements that shape daily economic activity.” This entails rules that establish and clarify property rights, moderate the cost of resolving disputes, boost predictability of economic exchanges and guard contractual partners against abuse by public or private agents.

nation. Developing nations that cannot commit to the whole package are unlikely to be able to attract the offshored factories and are thus unlikely to see their supply chain trade take-off.

Figure 2. Disciplines underpinning international production sharing



Source: Author's elaboration of diagram in Baldwin (2011a).

Not just RTAs: 21st century regionalism

The rise of international supply chains between high-technology developed nations and low-wage developing nations created new demand for and supply of international disciplines.

The demand came from advanced nations (and their firms) seeking to increase their competitiveness by offshoring certain stages of production. The supply came from developing nations, many of which opted to remove 21st century trade barriers to attract offshored factories and jobs. Joining international supply chains became the fast lane to industrialisation and growth, at least in nations near high-technology offshorers (e.g. the United States, Germany and Japan). Given this mutual interest in promoting international production sharing, the governments of developing nations willingly embraced disciplines on aspects of trade that were not traditionally considered as barriers. Specifically, the deeper discipline arose through three main policy “vehicles”: deep RTAs, BITs and unilateral reforms. We consider these in reverse order.

Unilateral reforms by developing nations

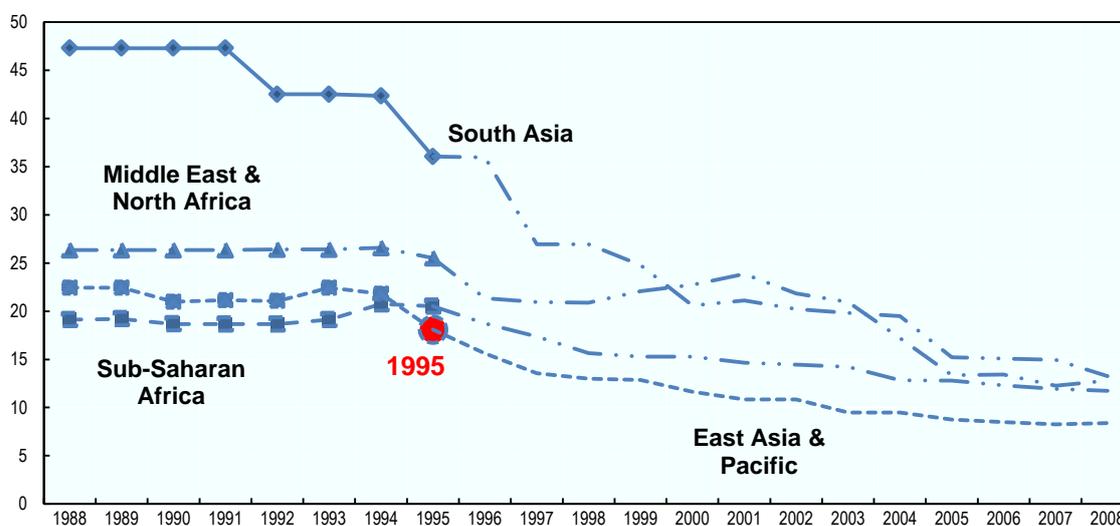
In the late 1980s and early 1990s, many developing nations engaged in full-throttle unilateral tariff cutting, evidenced in spectacular tariff reductions (Figure 3). Though this was partly driven by International Monetary Fund (IMF) conditionality (especially in Africa), even nations not subjected to external pressure lowered their rates. According to new evidence (WTO, 2011), the global tariff reduction on parts and components exceeds the overall average, providing rough evidence of an association between the second unbundling and an autonomous tariff liberalisation.⁷

The number of developing nations signing BITs also exploded between 1985 and 1995 (Figure 3). In essence, BITs provide unilateral concessions to rich-nation firms seeking to invest in developing nations (Egger and Merloz, 2012; Berger, 2008), i.e. they establish disciplines that govern interactions between private foreign investors and host governments. As such, they are central to the trade-investment-services nexus at the core of international production sharing, i.e. 21st century trade.

7. On the political economy of unilateralism, see Garnaut (1991), Young (1996), Edwards and Lederman (1998), Richardson (2001), Sally (2008), Coates and Ludema (2001), Krishna and Mitra (2008) and more recently, Ludema, Mayda and Mishra (2010), Conconi and Perroni (2010) and Baldwin (2010).

Figure 3. Unilateral tariff liberalisation, 1988 – 2008

Average developing nation tariffs by region



Source: Data from World Databank, World Bank.

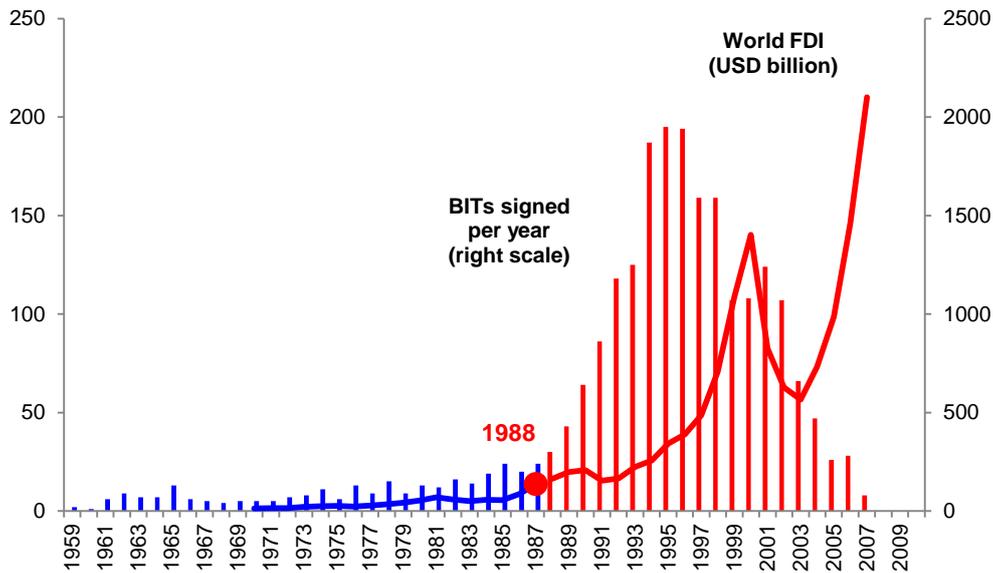
BITs are extremely common; about 2 800 exist worldwide. All the major foreign direct investment (FDI) emitters – Europe, the United States, Japan, etc. – have their own model agreements. The US model, which is quite explicit and comprehensive, illustrates well the basic features of a BIT.⁸

The basic goals of a typical US BIT are: (i) to assure non-discrimination in national treatment (in other words, US investors should be treated as favourably as national investors and third-country investors); (ii) to limit expropriations and ensure proper compensation when expropriations are unavoidable; (iii) to ensure that investors can move investment-related funds in or out of the country; (iv) to limit “performance requirements” placed by host nations on foreign investors (local content restrictions, etc.); and (v) to give foreign investors the right to submit disputes to international arbitration rather than local courts. The main arbitrator is the International Centre for Settlement of Investment Disputes (ICSID).

As Figure 4 shows, the number of BITs exploded in the late 1980s and early 1990s, around the time that developing nations embraced unilateral tariff cutting. Before this period, BITs occurred mainly between European FDI emitters and developing nation FDI seekers; the United Kingdom and Japan signed their first treaties in the 1970s and the United States only in the 1980s. The list of signers expanded rapidly in the 1990s. Since then, many developing countries have signed BITs with the major FDI emitters (the larger European Union countries, the United States and Japan), as well as other developing nations in their region. By now, almost all WTO members have signed multiple BITs.

8. See, for example, <http://www.state.gov/documents/organization/117601.pdf>.

Figure 4. Take-off in BITs and FDI



Sources: BITs from ICSID; chart adapted from Baldwin and Lopez-Gonzales (2013).

Deep RTAs

At about the same time as unilateral tariff cutting and BIT signing came into favour, RTAs with “deep” provisions – where deep means disciplines that help underpin GVCs (e.g. assurances for IP, capital movements, competition policy, business visas, etc.) – increased massively.⁹

Systematic data on these provisions first appeared in the form of a database (WTO, 2011) founded on seminal work by Horn, Mavroidis and Sapir (2010). The three authors read through all US and EU agreements, noting whether they contained:

- “deeper-than-WTO commitments”, i.e. commitments on areas already covered by WTO agreements, but where the RTA parties went deeper;
- “beyond-WTO-commitments”, i.e. disciplines on areas not covered in WTO agreements (e.g. free movement of capital linked to FDI).¹⁰

The authors also noted whether these provisions were legally enforceable.¹¹

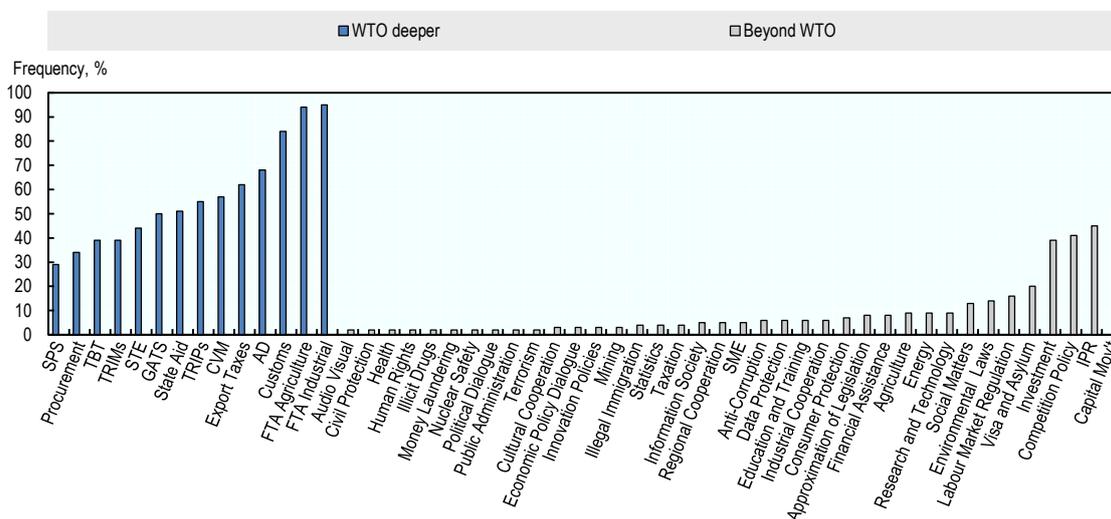
9. Lawrence (1996) highlighted explicitly the distinction between deep and shallow RTAs by. He also noted its association with more complex trade and pointed out that it first developed among developed nations in Europe and North America. Leaving aside the European Union’s Single Market –the ultimate deep RTA – the trend in deep RTAs started with the US-Mexico component of the North American Free Trade Agreement (NAFTA) and Europe’s Euro-Mediterranean Association Agreements (see, for example, Hufbauer and Schott, 2005 and 1993). Japan joined the movement by signing deep Economic Partnership Agreements (EPAs) with its large ASEAN offshoring partners (see Bilboa, 2008).

10. Horn, Mavroidis and Sapir (2010) call deeper-than-WTO commitments “WTO+” and beyond-WTO commitments “WTOx” provisions.

11. WTO+ provisions concern commitments that already exist in WTO agreements, but go beyond the WTO disciplines. WTOx provisions cover obligations that are outside the current WTO aegis. Yap, Medalla and Aldaba (2006) and Balboa (2008) did a similar exercise on Japanese EPAs.

The WTO database applies the same methodology to 100 RTAs. It provides information on 52 categories of provisions. Many of these are highly idiosyncratic, since European Union RTAs contain numerous issues that are only tangentially related to trade.¹² Figure 5 shows how often the agreements include the provisions in legally binding language.

Figure 5. Frequency of legally enforceable “deeper-than-WTO” and “beyond-WTO” provisions



Source: WTO RTA database, 2011.

Table 2 shows provisions that provide disciplines for 21st century trade. Some clearly aim to protect the tangible and intangible assets (e.g. intellectual property rights (IPR), capital movement and investment provisions) of foreign firms that offshore production to a developing nation. Others aim to better connect production facilities. For example, many General Agreement on Trade in Services (GATS) commitments involve liberalising infrastructure services (telecoms, express mail, air cargo, etc.). While these do not exclusively serve international supply chains, they do provide a critical element in the pro-supply chain package of disciplines.

To study the increase in deep RTAs in the late 1980s and early 1990s, we focus on the provisions most plausibly linked to international production sharing, namely: the Agreement on Trade-Related Investment Measures (TRIMs), the Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPs), competition policy, IPR, investment, movement of capital, approximation of legislation, industrial tariffs, customs and GATS. Figure 6 plots the total number of such provisions included in the stock of RTAs signed between 1958 and 2011. It also plots the total stock of all RTAs signed (i.e. those in the WTO data base). The results illustrate the sharp acceleration in deep provisions in RTAs, in conjunction with the boom in unilateral tariff cutting and BIT signing.

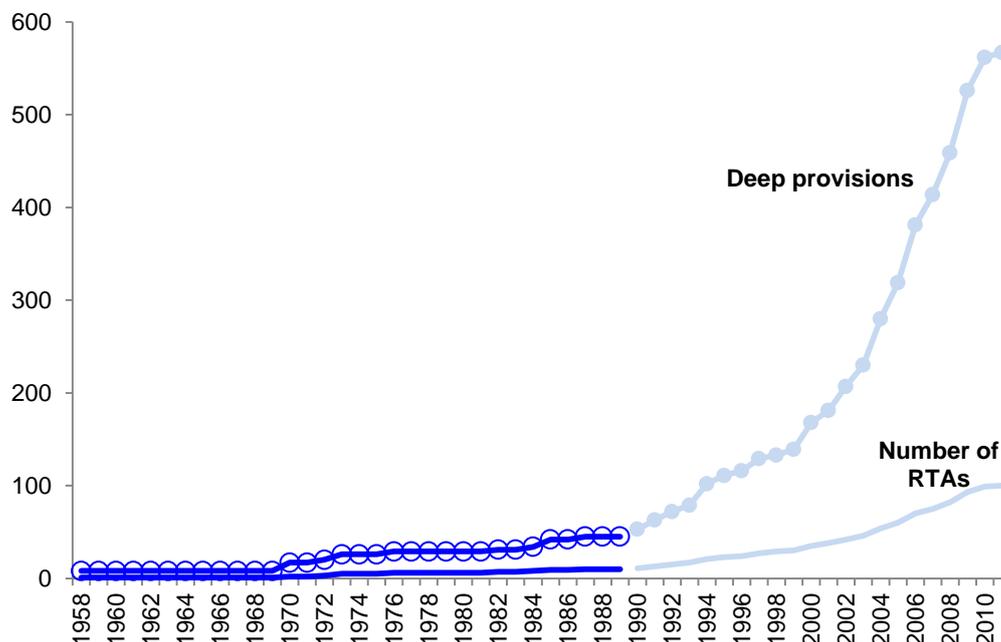
12. The European Union frequently uses RTAs as a form of foreign policy, so many non-trade issues creep in.

Table 2. Example of deep RTA provisions in the WTO database

Customs	Provision of information; publication on the Internet of new laws and regulations; training.
State trading firms	Establishment or maintenance of an independent competition authority; non-discrimination regarding production and marketing condition; provision of information; affirmation of Article XVII, GATT provision.
State aid	Assessment of anti-competitive behaviour; annual reporting on the value and distribution of state aid given; provision of information.
Public procurement	Progressive liberalisation; national treatment and/or non-discrimination principle; publication of laws and regulations on the Internet; specification of public procurement regime.
TRIMs	Provisions concerning requirements for local content and export performance of FDI.
GATS	Liberalisation of trade in services.
TRIPs	Harmonisation of standards; enforcement; national treatment, MFN treatment.
Competition policy	Maintenance of measures to proscribe anti-competitive business conduct; harmonisation of competition laws; establishment or maintenance of an independent competition authority.
IPR	Accession to international treaties not referenced in the TRIPs Agreement.
Investment	Information exchange; development of legal frameworks; harmonisation and simplification of procedures; national treatment; establishment of mechanism for the settlement of disputes.
Capital movement	Liberalisation of capital movement; prohibition of new restrictions.

Source: WTO www.wto.org/english/res_e/publications_e/wtr11_dataset_e.htm.

Figure 6. Deep RTA provisions and number of RTAs



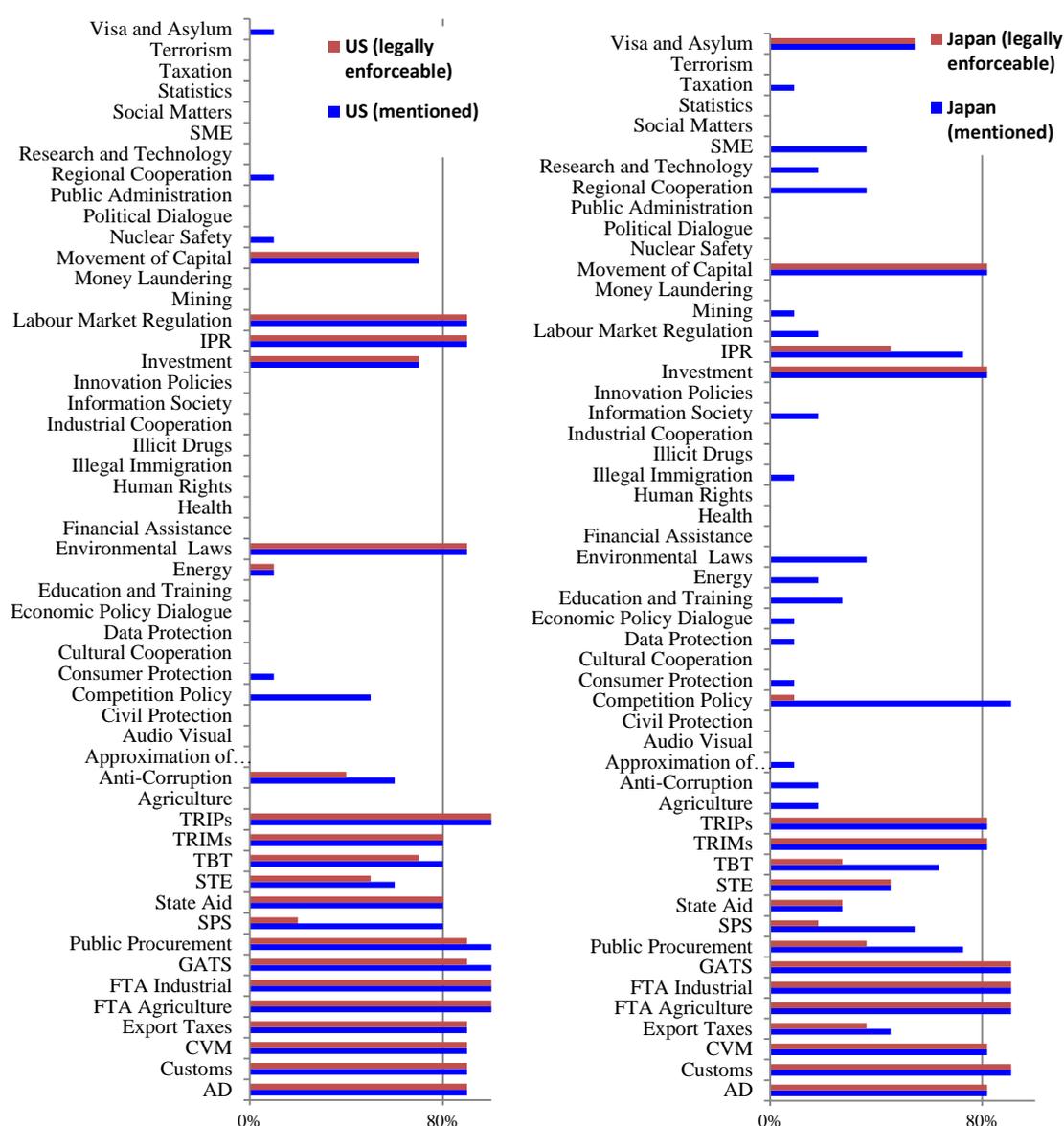
Source: Data from www.wto.org/english/res_e/publications_e/wtr11_dataset_e.htm.

US and Japanese deep RTA “templates”

As it turns out, the deep provisions included in RTAs by two of the largest organisers of international supply chains – the United States and Japan – feature some very stark patterns. Figure 7 shows the frequency with which each of the provisions shows up in US (left panel) and Japanese (right panel) RTAs included in the WTO database. The blue bars show the share of all agreements that mention each provision; the red bars show the share of provisions with legally enforceable language.

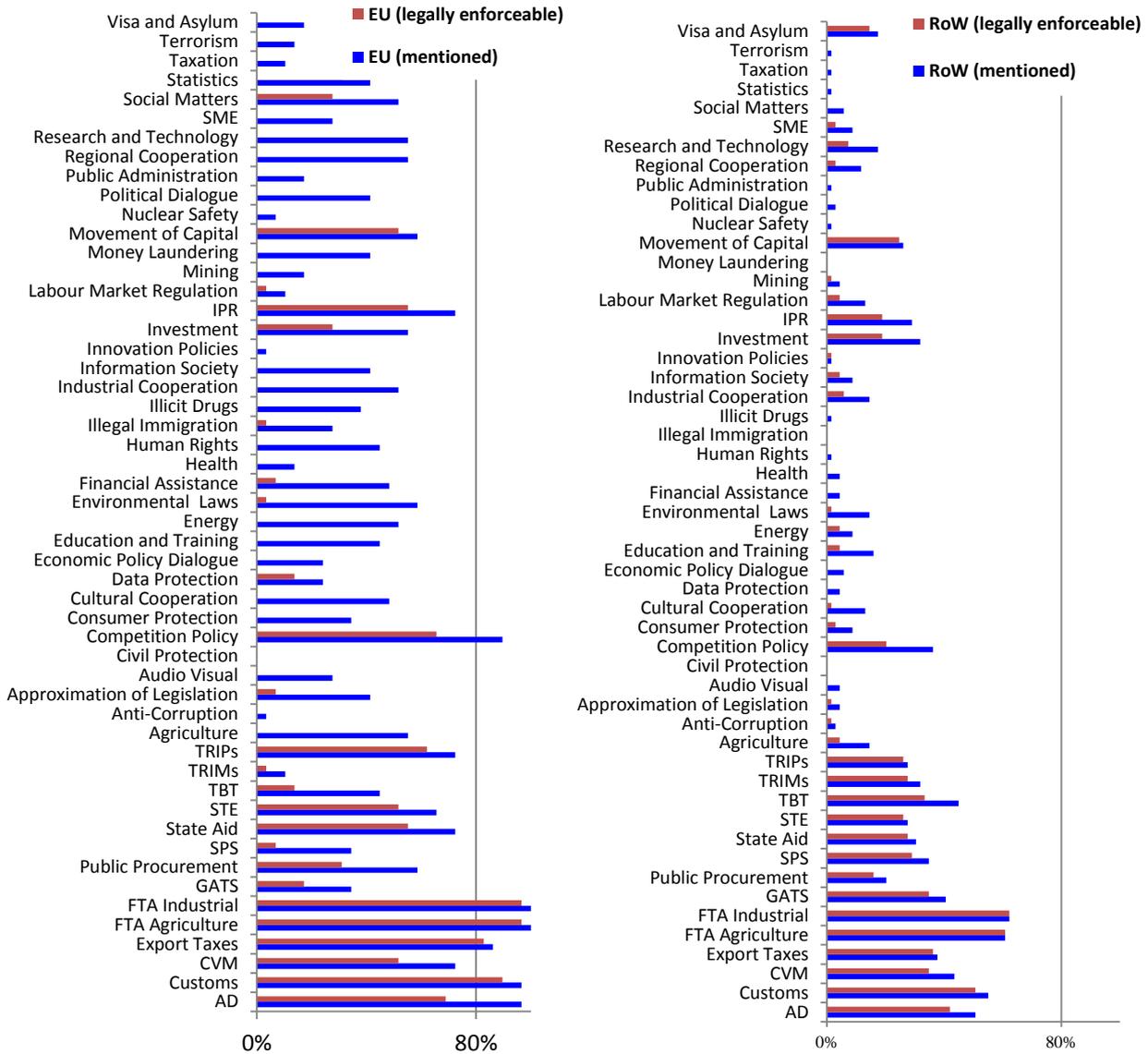
As we see, the United States is remarkably consistent in its provision coverage. Only 12 of the 52 provisions are included in most US RTAs, taking 80% as the threshold for “most”. While most of the 12 provisions involve deeper-than-WTO disciplines, they also include beyond-WTO disciplines in services, TRIPs, TRIMs, customs co-operation and procurement. These aim to encourage international supply chains (particularly IPR, investment restrictions and assurances) and the free movement of capital. Japanese RTAs show a fairly similar basic pattern.

Figure 7. Share of US and Japanese agreements with deeper provisions



Source: WTO RTA database, 2011. Adapted from Baldwin (2012b).

Figure 8. Share of European Union and rest of world (RoW) agreements with deeper provisions



Source: WTO RTA database, 2011. Adapted from Baldwin (2012b).

As we noted above, the European Union’s RTAs are much more diverse (Figure 8, left panel). Most EU firms have international supply chains within the European Union itself. Hence, the disciplines for German supply chains in Poland (for example) are underpinned by the ultimate deep RTA – the Single Market.

RTAs that do not involve the world’s largest organisers of international supply chains – namely, the United States, Japan and the European Union – are very shallow and only rarely include 21st century disciplines. This shows that 21st century RTAs are really about North-South supply chains.

3. Economics of 20th century and 21st century trade

Globalisation is often mistakenly viewed as linear – a progressive integration of national economies driven by lower technical and man-made trade costs. The traditional analysis of regionalism is based on this misunderstanding (Krishna, 2013; Bhagwati, 2008). In reality, globalisation leapt forward with two breakthroughs in “connective” technology, namely, transportation and transmission (Baldwin, 2006b).

The transportation breakthrough is called the “first unbundling”. It allowed consumption and production to be separated by great distances, while production stages remained bundled in factories and industrial districts. Its main impact was the increased ease with which goods crossed borders.

The transmission breakthrough (the information and communications technology (ICT) revolution) is the “second unbundling”. It allowed production stages to be unbundled and dispersed across international boundaries. Its main impact was felt in the increased ease with which ideas crossed borders.

The two unbundlings are very different, a point that is not widely understood. This section lays out the basic economic differences using simple partial equilibrium diagrams.¹³ We first touch on the essential differences between 20th and 21st century trade.

20th and 21st century comparative advantage

Traditional (20th century) trade concerns goods where the producing nation accounts for the vast majority of the export’s value added. In this case, exports are basically a bundle of national technology and production factors. Comparative advantage is purely a national concept.

The liberalisation of 20th century trade allowed nations to exploit their comparative advantage better by trading more, focusing production on what they did best while importing the rest. As productive efficiency grew, higher trade went hand in hand with higher welfare. Rather than alter nations’ comparative advantage, 20th century trade agreements strengthen existing comparative advantages. To put it differently, the trade system is used to sell and buy goods.

Twenty-first century (supply chain) trade concerns the complex international flows of goods, services, ideas, capital and people that arise when production processes are internationalised. While trade of goods is the most easily observed and best measured of these flows, trade is not the heart of the matter. The key to 21st century trade is the recombination of technology and factors across nations. In its most direct form, 21st century trade involves high-tech firms from high-wage nations that combine their managerial, marketing and technical know-how with low-wage labour in developing nations. This “technology lending” bears many names: foreign affiliates, joint ventures, contract manufacturing, offshoring, re-importing, export platforms, etc.

A more indirect form of 21st century trade involves importing intermediates that embody foreign technology and productive factors. Here, technology and factors are recombined across nations through the transfer of foreign know-how and factors embodied in imported parts and components (Jones, 1980; Deardorff, 2005).

In both cases, comparative advantage becomes a multinational concept. A key point here is that the trade system is being used to make things. Consequently, 21st century trade agreements can alter comparative advantages. One nation’s exports become competitive based not so much on the easier movement of goods, but on the easier cross-border movement and combination of several nations’ technology, labour and capital in the context of internationalised production networks.

13. For example, basic Vinerian economics assume goods derive 100% from the exporting nations and ignore the impact of RTAs on production sharing. This 20th century view still dominates thinking on RTAs (see Krishna, 2013).

Twentieth-century trade: Basic economics

To contrast 20th century and 21st century regionalism, we start with the familiar Vinerian analysis.

The well-known RTA diagram (Figure 9) is a three-nation (“Home”, “Partner” and “RoW”) framework with three goods, numbered 1, 2, and 3.¹⁴ To enable comparison, we assume all nations are symmetric in size and tariff levels. All markets are also symmetric, but comparative advantage is staggered – each nation exports two goods and imports the other. Since each nation has two sources of imports, each market can display tariff discrimination.

We study the market for only one good, the good that Home imports. The diagram shows the relevant export supply curves (marked XS) for Home’s two potential suppliers (Partner and RoW). The horizontal sum of the XS curves, in the absence of tariffs, is MS^{FT} . All nations have a specific “T” tariff on all goods to start with, so the sum of import supply curves is MS^{MFN} (MS^{FT} shifted up by T). Home’s internal price is “P”; the border price is $P - T$ (which is what Partner and RoW exporters receive).

When Home and Partner sign an RTA, tariffs between them drop to zero. MS shifts to MS^{FTA} and Home’s internal price falls to P' . There are two border price effects. Partner sees its export price rise to P' , while RoW sees its export price fall to $P'-T$. Partner exports expand (trade creation) and RoW exports contract (trade diversion). Identical things happen in the market for the good where Home is the exporter and Partner and RoW are the importers.

Three elemental Vinerian effects

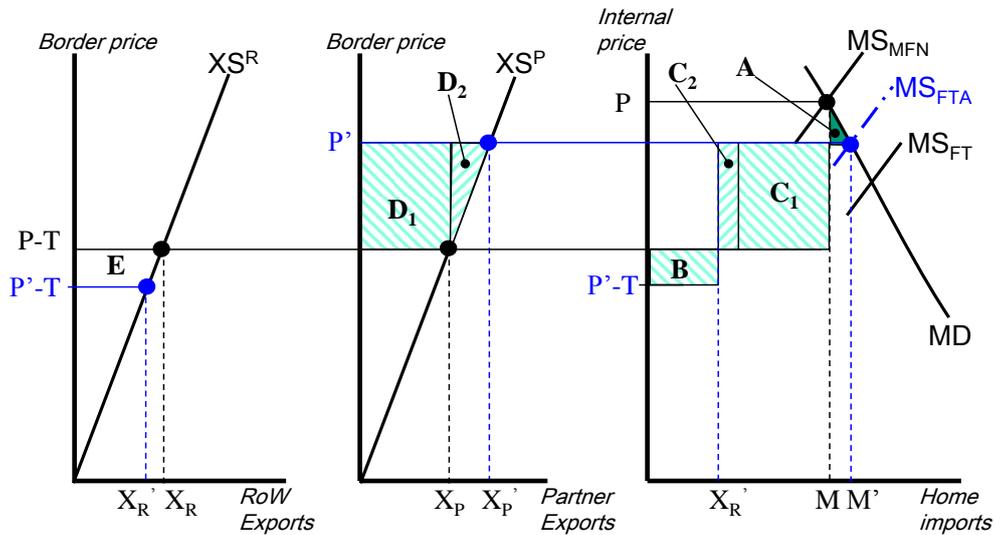
In terms of international political economics, three elemental effects apply (Baldwin, 2009). Smith’s Certitude notes that tariff preferences benefit exporters who receive them (area $D_1 + D_2$ in the diagram). Haberler’s Spillover notes that tariff preferences harm exporters who do not receive them (area E). Viner’s Ambiguity points out that discriminatory tariff removal has ambiguous overall welfare effects on the RTA partners.¹⁵

The domestic political economics are equally simple. Domestic trade usually turns on commercial interests. Consumers are usually disorganised and governments (at least in rich nations) tend to underplay tariff-revenue considerations when considering trade deals. Home’s import-competing producers oppose the RTA because it lowers the price to P' . Home producers of its export support it because their export price rises to P' (remember that all markets and nations are symmetric in this illustration). Thus, the politics of the RTA pits Home’s exporters against Home’s import competitors. A similar array of special interests determines Partner’s stance on the deal.

14. Many authors use the small-open economy version of this diagram, which would prevent us from considering third-nation effects that are the heart and soul of the Vinerian comparison of multilateral and regional tariff cutting. The small-open version is useful for considering the impact on the home country in isolation.

15. Home welfare impact in the good Home imports is $A+B-C_1-C_2$, but Home gains the equivalent of D_1 and D_2 in the market for the good it exports to Partner. Since $D_1 = C_1$, the net effect is $A+B+D_2-C_2$.

Figure 9. The RTA diagram



Source: Author's elaboration on diagrams in Baldwin and Wyplosz (2012, Chapter 5).

Non-tariff barriers, slanted multilateralism and negative trade diversion

While classroom analysis of regionalism almost always focuses on tariffs, recent RTAs often include deeper provisions, as discussed above. Since the differences between tariff and non-tariff barriers have important implications for interpreting the empirical evidence, we illustrate this trend with the PTA diagram.

Trade falters when the price of imports in the importing nation differs from the price in the exporting nation. Free trade is defined as the absence of such a gap. The gap comprises many elements, such as tariffs, logistics (e.g. shipping, insurance, port clearance fees, etc.) and technical barriers to trade (TBTs) – the cost of adapting foreign products to the importing nation’s particular regulatory regime. Figure 9 assumes the whole gap was due to a tariff. In the next illustration, the whole gap stems from a “frictional” barrier, which does not generate rent or revenue to anyone as it merely raises the cost of importing.

We can alter the basic PTA diagram to determine how the reduction of regulatory and related barriers (TBTs, poor infrastructure services, bad ports, etc.) can be so different. To keep things simple, we suppose the tariff-equivalent of the frictional barriers is “T”. We study two policy experiments: fully discriminatory and partially discriminatory frictional barrier liberalisation.

In the first experiment, Home completely eliminates T for imports from Partner, but not for imports from RoW. In the second experiment, eliminating T in the context of an RTA with Partner has a positive spillover for RoW, i.e. the Home-Partner RTA leads to lower frictional barriers for RoW, but T falls to T’ rather than zero. The point here is to reflect the reality that since many NTB liberalisations come without rules of origin, the beneficial effects tend to be less exclusive (as argued above).

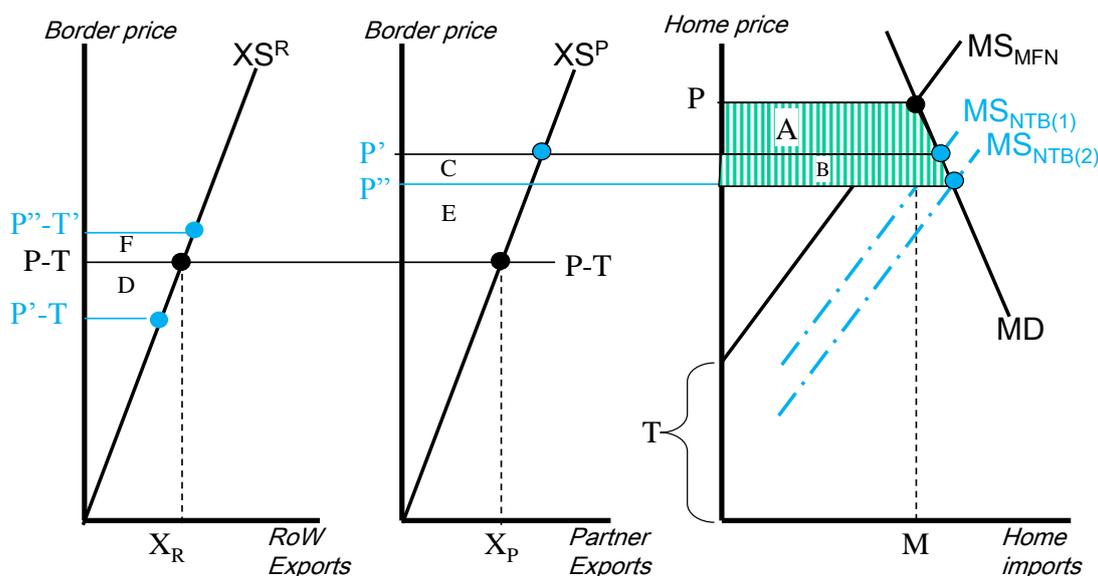
Figure 10 shows the analysis. Again, the initial situation is where T is applied to imports from Partner and RoW, so Home’s MS is MS_{MFN} . The **fully discriminatory liberalisation** has positive effects that are identical to the discriminatory tariff liberalisation. The relevant MS curve is $MS_{NTB(1)}$. The new price is P’, which means a rise in the export price for Partner, but a drop in the export price of RoW. These price effects produce the usual trade creation and trade diversion. Home gains area A, Partner gains area C+E and RoW loses area D. Since there is no loss in tariff revenue, Viner’s ambiguity disappears for frictional barrier liberalisation. Apart from the welfare effects on Home, this is identical to the Figure 9 analysis.

The **partially discriminatory liberalisation** has quite different trade and welfare effects. When the RTA between Home and Partner results in the full elimination of T for Partner exports and a partial reduction of T to T’ for RoW, the relevant MS curve is $MS_{NTB(2)}$. The Home price is lower at P’, so Partner sees its price rise to P’ and RoW sees its price rise to P’-T’. Note that both exporters see their price rise, but the rise is less important

for RoW. As usual, Partner exports more, so we should see trade creation. The unusual outcome is that we should also see “negative trade diversion” – RoW exports would also rise with the “preferential” agreement, but less than Partner’s exports.

The partially discriminatory liberalisation raises welfare for all three nations: for Home by A+B, for Partner by E and for RoW by F.

Figure 10. Preferential frictional barrier liberalisation



Source: Author’s own elaboration.

Summary

For traditional trade, preferential tariff cutting allows each nation to exploit its comparative advantage better by exporting more (trade creation). Since preference for one is discrimination for the others, this trade creation partly occurs at the expense of third-nation exports (trade diversion). For frictional barriers liberalised by RTAs, the trade effects may be similar to those of a preferential tariff cut. However, we may also see negative trade diversion, where the RTA boosts Home’s imports from both Partner and RoW, but RoW exports rise less than Partner’s. Section 4 below discusses the many studies that find evidence for “negative”, or “reverse”, trade diversion.

Twenty-first century trade: Basic economics

The PTA diagram does not allow us to consider how production internationalisation can shift trade patterns independently of trade liberalisation. It is therefore not the right framework to study 21st century regionalism and multilateralisation. Two new diagrams will serve this purpose.

This sub-section assumes that all trade in goods is costless and free of all natural and man-made barriers. Since in this context a 20th century RTA would have no effect whatsoever, any observed trade effects stem from factors unrelated to 20th century regionalism.

Twenty-first century RTAs foster two individual aspects of production unbundling: (i) direct recombination of national comparative advantages via foreign affiliates (e.g. offshoring); and (ii) indirect recombination via new trade in parts and components (i.e. the foreign technology and factors are embodied in imported intermediates). We illustrate each of these in turn.

Directly recombining Northern technology with Southern labour

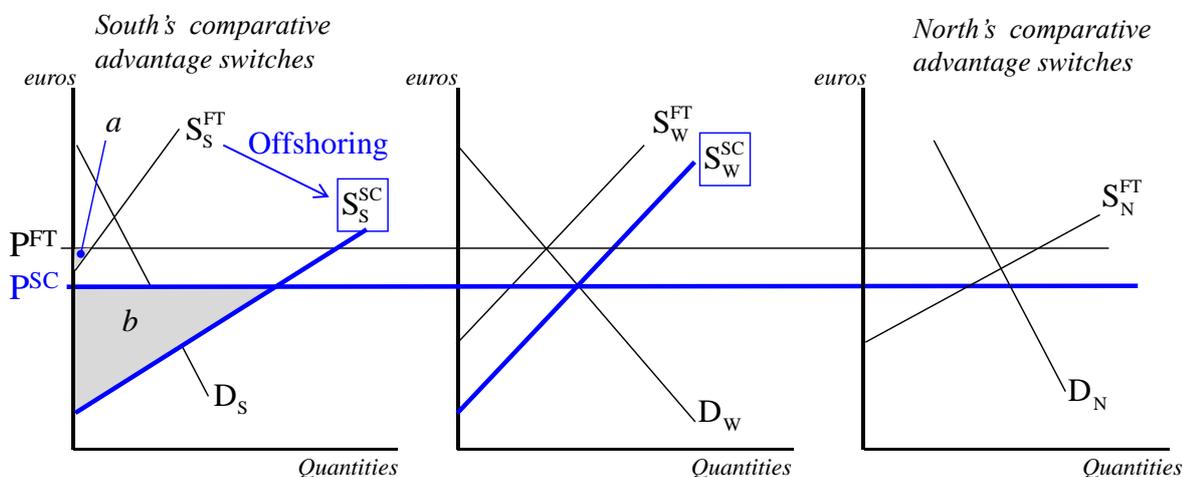
Let us consider a world with two nations where one nation has better technology, but higher wages than another nation. For convenience, call the high-tech, high-wage nation “North”, and the low-tech, low-wage nation “South”. To start from the familiar setting of 20th century trade, we open the analysis assuming free trade in goods, but no international mobility of technology. The point is to consider the impact of a deep RTA that alters South’s domestic policy environment in a way that makes it safe for North’s high-tech firms to apply their know-how in the South. After the deep RTA, the firms’ technology can move internationally, but stays under their control.

The first point to keep in mind is that technology will not flow both ways. North’s firms will find it profitable to combine their high-tech know-how with South’s low-wage labour. South’s firms, by contrast, will never want to combine their low-tech know-how with North’s high wages.

To highlight how different this outcome is from the traditional view of trade liberalisation, consider a historical analogy, the 19th century mass migrations to the New World. In the late 1800s, there was a big gap between the land-labour ratio in the Old World and the New World that made European wages lower than New World wages. As land could not be moved, labour went to land. This movement of nations’ “sources of comparative advantage” had big effects on world trade patterns. In particular, the New World became a major exporter of land-intensive goods such as food and cotton (O’Rourke and Findlay, 2009).

In the 21st century, the big gap is in the technology-labour ratio. Twenty-first century globalisation is very much about technology (broadly defined to include marketing, management and technical know-how) moving to labour. This outcome has many monikers – offshoring, fragmentation, vertical specialisation, production unbundling, production sharing, GVCs, etc. We call it supply chain trade, or 21st century trade. Note the difference with 20th century trade, where all the sources of comparative advantage are immobile and the goods trade is the only way of exploiting comparative advantage.

Figure 11. Impact of 21st century RTA: Switching comparative advantage



Source: Author’s elaboration.

To illustrate the effects in a simple diagram, we focus on the market for a single good, namely the good that North initially exports (Figure 11). The middle panel shows the world supply and demand curves (S_w and D_w), with superscripts indicating the regime: FT for free trade, SC for supply chain trade. The left panel shows South’s supply and demand curves.

Note that before the deep RTA is signed, South’s supply curve is S_s^{FT} ; North’s supply curve is S_N^{FT} , reflecting the combinations of national technology and national factors (North produces in North and South produces in South). Here, we assume that North initially has a comparative advantage in the good (i.e. North

exports the good since its technological advantage outweighs its labour cost disadvantage; this is why S_S^{FT} is higher than S_N^{FT}). As there are no tariffs or trade barriers of any kind, P^{FT} is the price in both nations.

After the deep RTA is signed, North's firms have the assurances they need to move their tangible and intangible assets to South. Given the wage differences, this is a profitable move and an international supply chain arises. Since production in South involves Northern technology and Southern labour, South's supply curve shifts down to S_S^{SC} . Nothing happens to North's supply curve, since its production still uses its own technology and labour. In a nutshell, the main changes are:

- World supply shifts out to S_W^{SC} , the price drops from P^{FT} to P^{SC} and global production increases;
- Trade volume rises, but North and South swap comparative advantage;
- South switches from importing to exporting the good;
- North stops production and starts importing the good.

Note how this "trade creation" does not stem from lower trade barriers (trade is costless both before and after the RTA). Rather, the trade volume rises due a recombining of the nations' sources of comparative advantage.

The domestic political economics are very different for supply chain trade than they are for traditional trade. To see this, we focus on three groups in each nation – firms, workers and consumers (remember that there is no tariff revenue, so the government is not directly affected). To link production and workers' interest naturally, we assume that a nation's firms share a fixed share of their producer surplus with the local workers.¹⁶

The political economy follows from welfare effects on the groups. North's firms see their supply curve drop when they can use cheaper labour (far right panel). The producer surplus is area c under the FT and area b with supply chain trade, so North's firms gain c-b; if this is not positive, no offshoring arises. South's firms exit and lose area a, but Southern workers are likely to gain as there is more producer surplus to share. Thus, while North's firms and consumers favour the SC regime, North's workers oppose it. Meanwhile, South's firms oppose it, but workers and consumers support it.

Notice how standard commercial pressures for tariffs disappear with this type of liberalisation. Since North's firms are producing in South, they do not lobby for protection in North; South's firms no longer import, so tariffs become useless. If our examples included tariffs, all political economy support for the initial tariffs would evaporate after the offshoring. As for third nations, the price drop would harm the welfare of exporting nations and help that of importing nations.

Let us now examine how 21st century RTAs can switch comparative advantage more indirectly.

Unbundling upstream and downstream production

In the previous example, international supply chain arises only in final goods. In the real world, production unbundling typically involves intermediate goods (Ando and Kimura, 2005), so we introduce a diagram to capture this important aspect of 21st century trade.¹⁷ To spotlight the indirect recombination, we revert to assuming that technology is immobile, but continue to assume that trade in goods is perfectly free.

In the pre-RTA world, all production stages in both nations are bundled spatially (in factories) to save on communication and co-ordination (not trade, as those are set to zero) costs. Moreover, exports consist of 100% local value added.

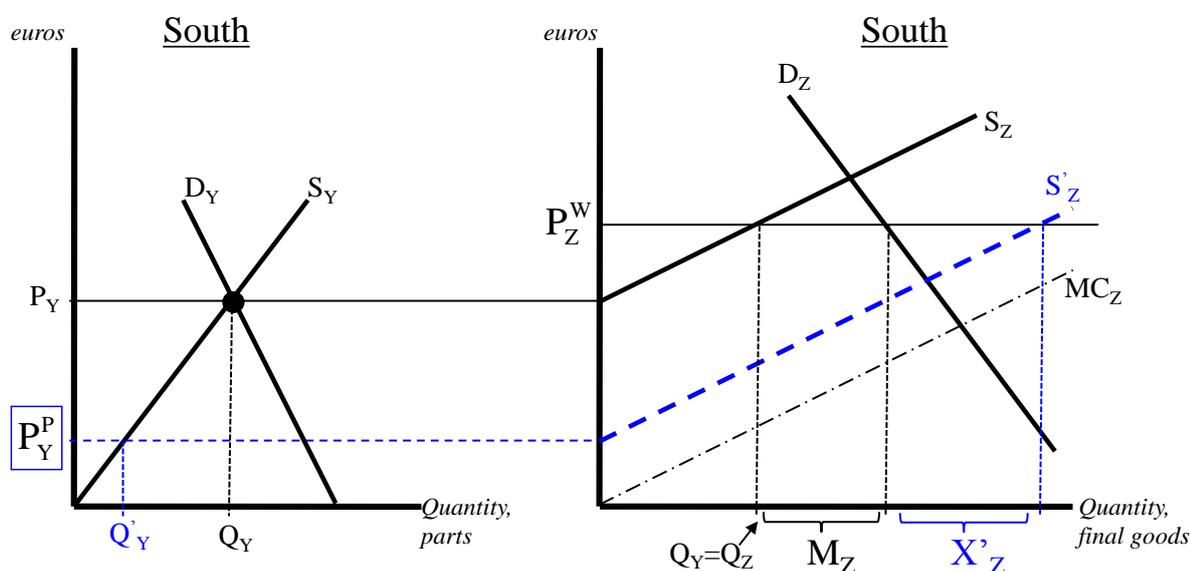
16. In a model where firms and trade unions have bargaining power, the equilibrium the share could be determined by Nash bargaining.

17. Also see Gereffi (2001) for early examples and the website <http://www.globalvaluechains.org> for abundant recent case studies.

Consider what happens when this initial situation is changed by two developments. The costs of co-ordination fall with the ICT revolution and the nations sign a deep RTA that provides the assurances needed to stimulate offshoring. As a result, production of intermediate goods and final goods take place in different nations. In other words, the production process is internationalised.

We illustrate this with a simple diagram where there a single upstream (i.e. intermediate) good called Y and a single downstream good called Z (the mnemonic is that Y comes before Z, so Y is upstream of Z). The linked diagrams (**Error! Reference source not found.**12) show the equilibrium in South for Y (left panel) and Z (right panel); S is the supply curve and D is the demand curve, while the subscripts indicate the good. We assume the input-output linkage to be as simple as possible; each unit of Z requires one unit of Y.¹⁸ In addition to the cost of input Y, there is a marginal local labour cost MS_Z to produce Z (right diagram). The supply curve for Z is the vertical sum of MC_Z and the price of Y.

Figure 12. Supply chain trade with intermediate goods and no technology lending



Source: Author's elaboration.

The idea here is that before the deep RTA and ICT revolution, it is too expensive or risky for South to produce Y and Z in separate nations, so South makes its own Y locally. When this is true, the supply curve for Z is MC_Z plus the equilibrium price of P_Y (set in the left panel). While South is producing Y and Z, given S_Z , South imports Z in an amount indicated as M_Z (right panel).

The RTA changes this. Its deep disciplines lead to reliable supply chain logistics (express mail, air cargo), telecommunications and business mobility (key managers and technicians moving to co-ordinate Y and Z production), etc. Consequently, South can import Y. To obtain sharp results, we assume that after the RTA, supply chain linkages become costless and perfectly reliable. Thus, South can start to import Y from the RTA Partner at the lower price of P_Y^P , reducing its own Y production to Q_Y^P . The lower price for Y lowers S_Z to S'_Z .

The key effects of this are that South switches from importing to exporting Z and starts to import Y. Note how this trade creation did not stem from trade liberalisation per se (as there were no trade barriers before), but rather from relaxed co-ordination (as opposed to transportation) constraints¹⁹.

18. D_Y is a derived demand curve, i.e. linked to the production of Z.

19. Obviously, one could find a tariff equivalent of the co-ordination barrier. Empirically, though, trade costs fell little after the second unbundling, while ICT costs plummeted.

Although we have not shown North (who produces the intermediates at a lower cost than South) explicitly, it would be easy to draw a case where it exports Z before the RTA, but imports Z and exports Y after the RTA. This is a clear example of how foreign technology and factors embodied in the imported component can transform South's comparative advantage, creating new trade for South in Y and reversing its pattern of trade in Z.

Using the same three groups as before – firms, workers and consumers – the impact on South is clear. Workers and firms in industry Y lose, as both the price and sales volume are lower. Workers in Z gain, as the combined surplus of X and Y is maximised by sourcing inputs from the lowest-cost source. This suggests that South's government will favour the deep RTA, which leads to an overall expansion of its producer surplus.

There is no mystery in this outcome. Before the RTA, South had a latent comparative advantage in Z, but a latent comparative disadvantage in Y. The RTA allows South to specialise in its comparative advantage sector.

While we have kept the RTA partner in the background, we can imagine several scenarios. The easiest is to assume that North was exporting Z to South pre-RTA and is exporting Y and importing Z post-RTA. In this case, Northworkers and firms in industry Z would oppose the RTA, while those in Y would support it.

4. The impact of regionalism: Empirics

The traditional 20th century view of the economic impact of RTAs centres on the amount of trade created and diverted. This view gained intellectual dominance among policymakers and economists in the 1950s and 1960s, when tariffs were high worldwide. But this dominance is not founded on empirical evidence. Decades of empirical testing have yielded few solid results. Most of the work is based on the view that RTAs are mainly about preferential tariff liberalisation and indeed, several recent high-profile publications assume they are *only* about tariffs (Romalis, 2007). Given the above discussion of the limited relevance of tariff preferences in the 21st century, much of this work is no longer relevant. Nevertheless, we discuss the studies for completeness.

Shallow RTAs: Trade creation and diversion

Clausing (2001) uses detailed tariff data to show that the Canada-United States FTA had substantial effects on trade creation, with little evidence of trade diversion. Fukao, Okubo and Stern (2003) find that North American Free Trade Agreement (NAFTA) tariff preferences had effects on trade creation and diversion in a handful of industries (e.g. clothing, cloth and some footwear products) where MFN tariffs applied by the United States remained high. However, the tariff preferences had no statistically important impact on most industries (e.g. motor cars and vehicles and television receivers). The authors take this as evidence that NAFTA was working via 21st century trade channels (i.e. outsourcing) rather than tariff preferences.

More recently, Freund and Ornelas (2010) note that the conventional empirical strategy involves estimating trade creation and diversion via the gravity equation, allowing researchers to control for income and idiosyncratic factors related to the year, bilateral relationship and individual nation. They state that while “the empirical literature is not entirely conclusive, it does suggest that trade diversion is not a major concern, though in some agreements and sectors it may matter.”

One important study in this line, Magee (2008), uses data from the late 20th century to estimate the creation/diversion effects of 15 separate RTAs, including NAFTA, the 1986 enlargement of the European Union, the 1992 bilateral RTAs between the European Union and several Central European nations, Mercosur, ASEAN and select RTAs among Latin American as well as African nations. Only 8 of the 15 RTAs were found effective (controlling for other factors) in creating new trade among the partners.²⁰ The other RTAs neither increased nor decreased trade significantly. Moreover, Magee finds that NAFTA (the second-largest RTA, whose internal trade accounts for about 8% of global trade) created trade and produced “reverse trade diversion”. In other words, controlling for other factors, NAFTA made it easier, not harder, for excluded nations to export to the United States, Canada and Mexico. This strongly suggests it is misguided to apply a Vinerian

20. The eight are the 1986 European Union enlargement, the Andean Community, Mercosur, the ASEAN FTA, NAFTA and the Association Agreements linking the European Union and EFTA with Central and Eastern European countries.

analysis to the economic effects of NAFTA, as its provisions do not appear to act as preferential tariffs in the aggregate.

Egger, Larch, Staub and Winkelmann (2011) find the trade creation effect is substantially larger when using an econometric methodology that controls for reverse causality (i.e. the fact that nations that trade unusually much also tend to sign RTAs, so the high level of trade is causing the RTA and vice versa). Studying data on six Latin American RTAs, Freund (2010) shows that RTAs are not associated with trade diversion and that preferential tariff reduction tends to precede external MFN tariff reduction in a given sector. This leads her to conjecture that the lack of trade diversion may arise from unilateral MFN tariff cutting, which lowers tariff preferences.

On the other hand, studies of areas and time periods where MFN tariffs were still high tend to find more evidence for trade diversion. Chang and Winters (2002) examine export price data for five non-member countries' exports to Mercosur. They find that tariff preferences granted to Argentina, Uruguay and Paraguay caused substantial reductions in the prices received by United States, Chilean, German, Korean and Japanese exporters to Brazil.

Yeats (1997) finds that goods for which Mercosur members did not have a revealed comparative advantage displayed the fastest intra-Mercosur trade growth. They interpreted this correlation as evidence that Mercosur diverted trade in the late 1980s and early 1990s. Note that this was before Latin America's big unilateral MFN tariff cuts were fully implemented (see Figure 3), so Mercosur initially created large margins of preferences. Yeats' findings are probably no longer relevant as the unilateral tariff cuts by Brazil would have greatly reduced the Mercosur-linked tariffs preferences.

Deep RTAs and supply chain trade

Analysts have only recently begun to focus on the impact of deep RTAs on trade through other channels than trade creation and diversion. The seminal analysis by Orefice and Rocha (2013) shows that deep RTAs tend to foster trade in production networks among RTA members (with an average effect of 12 percentage points). Their strongest findings are for industries where international production networks are most prevalent, e.g. electrical and mechanical machinery. For example, while deeper RTAs boost supply chain trade by 36% in automotive parts and 11% in information technology (IT) products, the impact is statistically insignificant for trade in textiles. The result for IT products, on the other hand, is particularly significant as tariffs in this industry have been set to zero at the multilateral level since 1997.²¹ This is highly informative as we know there can be no tariff preferences in this industry, so the trade creation result must stem from the deeper elements of the RTAs.

The authors also find that two nations that already trade in production networks are more likely to sign a deep (as opposed to a shallow, tariffs-only) RTA. The effect is five times stronger for North-South agreements. This latter finding is particularly interesting, since it is likely that RTAs are allowing Southern nations with weak institutions to use the RTAs to make their domestic pro-offshoring reforms more credible. Finally, the positive effect of production network trade on deep integration is driven largely by Asia, where international production sharing has flourished the fastest.

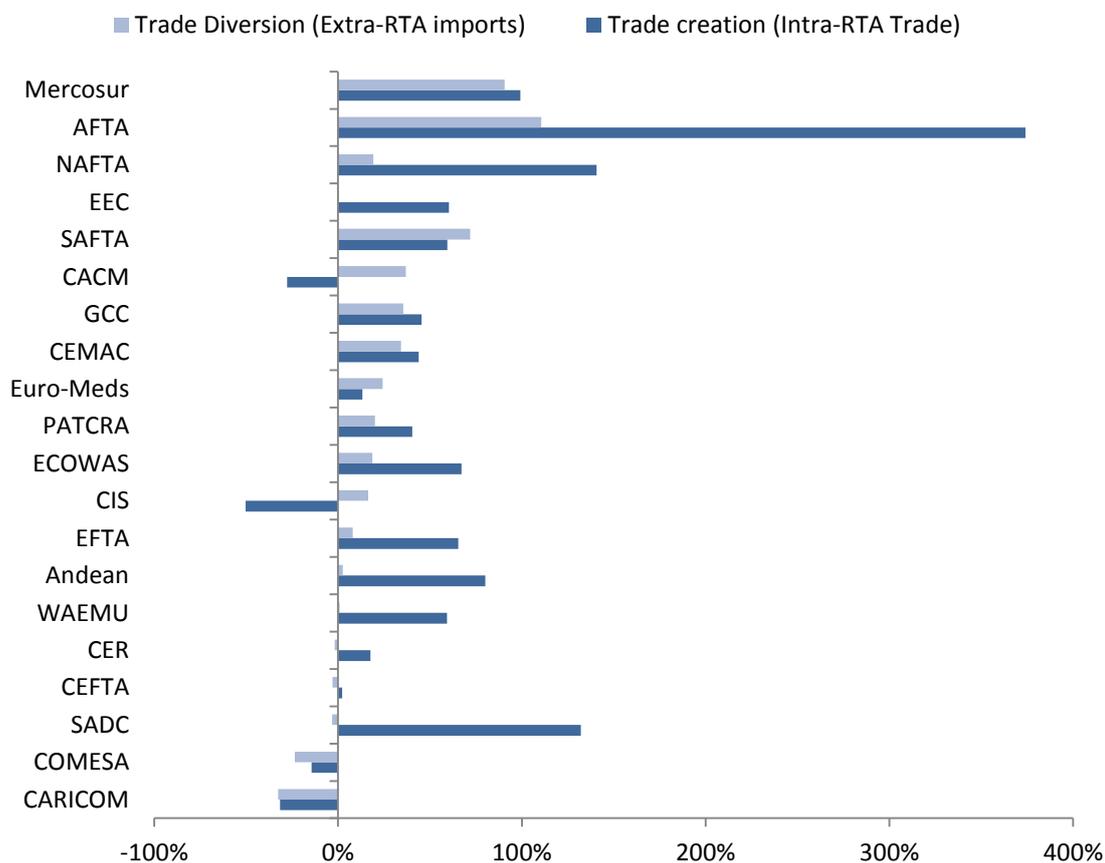
These findings help to explain the seemingly paradoxical rise of deep integration at a time when preference margins are shrinking. The conclusion is that partners are not primarily exchanging market access when they sign RTAs. Rather, they are locking in disciplines that foster offshoring and the 21st century trade that comes with it.

Damuri (2013) pursues a related empirical strategy, but uses a different measure of RTA deepness. He tests the hypothesis that countries that are intensely involved in international production networks are more likely to sign deep RTAs, especially with their most important production-sharing partners. He also decomposes the link between production sharing and deep RTAs into non-bilateral (or general) and bilateral (partner-specific) effects. The general effect – the extent to which a particular country wishes to be connected to an international production network – is captured by multilateral trade in parts and components. The partner-

21. This was accomplished through the Information Technology Agreement.

specific effect – the importance of a given partner in the country’s production network – is embodied in bilateral trade of parts and components. While deeper RTAs reflect both factors, the general effect dominates. This finding is quite robust to changing samples and specifications or the use of alternative indicators. The idea is that international production sharing is very much a network of related firms, rather than a simple bilateral connection between firms in two nations.

Figure 13. Recent estimates of trade creation and trade diversion



Source: Acharya et al. (2011).

Acharya et al (2011) present complementary findings, based on data drawn from the period after the massive unilateral tariff liberalisation documented in Figure 3. If one thinks of RTAs as being mostly about tariff preferences, their estimates show a curious pattern of trade creation and diversion (Figure 3). However, realising that tariff preferences were low for their estimation sample (see the related reasoning in Section 2), the findings are easily understood.

The most striking finding is that almost all of the RTAs have led to reverse trade diversion, i.e. external trade creation. In a 20th century analytic framework, this would make no sense as tariff preference for partners is tariff discrimination for third nations. Here, the RTAs create trade for member and non-members alike.

The most likely explanation is: i) the RTAs they look at liberalised frictional barriers that lower trade barriers for all nations, but more so for RTA members. As the Figure 10 reasoning showed, we should expect this to lead to trade creation among RTA partners and “negative trade diversion” with non-RTA nations. This suggests that RTAs act more like general trade liberalisation schemes – but slanted toward members – although where the Common Market for Eastern and Southern Africa and Caribbean Community are concerned, they seem to act as general trade restriction schemes.

Antras and Foley (2011) take a different approach, analysing the impact of the ASEAN FTA (AFTA) on the level and nature of US multinational firms' activity. They find that AFTA boosted the number of US firms investing in AFTA members, as well as the size and sales of the affiliates within AFTA markets. While this does not indicate a direct connection between 21st century RTAs and supply chain trade, it does illustrate that RTAs like AFTA can have effects far beyond the traditional trade creation and diversion framework presented in Figure 9. In this case, it seems that AFTA is affecting FDI patterns.

Summary

According to empirical evidence, tariff preferences no longer dominate regionalism in the 21st century. Studies of recent data find that RTAs lead to modest trade creation and reverse trade diversion, suggesting that traditional thought patterns (i.e. standard Vinerian analysis involving trade creation and diversion, as illustrated in Figure 9) are incorrect or incomplete. The most likely cause of this very non-traditional outcome is that combinations of GATT negotiations and unilateral liberalisations by developing nations have greatly reduced the importance of tariff preferences. Thus, we should stop thinking of RTAs in terms of tariff preferences and focus instead on their impact on frictional barriers and disciplines related to international production networks. The few papers that directly test the impact of 21st century trade agreements concur that deep RTAs tend to foster, and be fostered by, 21st century trade.

5. Liberalising 20th versus 21st century trade

As discussed above, 21st century regionalism is not a simple phenomenon. It consists of unilateral reforms by developing nations, BITs and deep RTAs. By contrast, 20th century regionalism was driven mainly a search for tariff preferences.

This section argues that the very different economics and political economy of 20th and 21st century trade account for the very different liberalisation vectors.

- Twentieth-century trade grew from a continuing process of tariff cutting through regionalism and multilateralism.
- Twenty-first-century trade, by contrast, did not progress at the multilateral level, but progressed rapidly at the regional level.

Since this contrast has critical implications for multilateralising 21st century regionalism, we briefly review the two phases of liberalisation and their associated political economy drivers.

Liberalising 20th century trade: Juggernauts and dominos

Modern multilateralism was born in 1947 with the signing of the GATT. Modern regionalism also started in 1947, when regionalism was woven into the GATT (Article 24) and the Organisation for EEC launched European regional integration.²²

In the early 1960s, regional liberalisation kick-started multilateral liberalisation. The rapid progress of the EEC customs union, combined with the United Kingdom's 1961 application to join, prompted the United States to seek to reduce preferences through the Dillon Round and Kennedy GATT Rounds (Ludow 2007; Dam 1970).

In 1964, regional and multilateral liberalisation boomed in tandem. The Kennedy Round began, the United States initiated talks on its first RTA (the United States-Canada Auto Pact), and Australia and

22. See Irwin, Mavroidis and Sykes (2008), Pomfret (1997, Chapter 4), Zeiler (1997), Dam (1970, Chapter 2), or Jackson (1997, Chapter 2.2). The United States had more direct interest in the Article 24 exception; in March 1948, the United States and Canada concluded a secret draft protocol eliminating most tariffs and quotas bilaterally, which was ultimately rejected by the Canadians for fear their industries would be crushed or absorbed by their US competitors (Smith, 1988, p.39; Wonnacott, 1987, p.15; Chase, 2007).

New Zealand signed an RTA, all in 1964. While South-South regionalism flourished on paper, few tariffs were actually lowered.²³

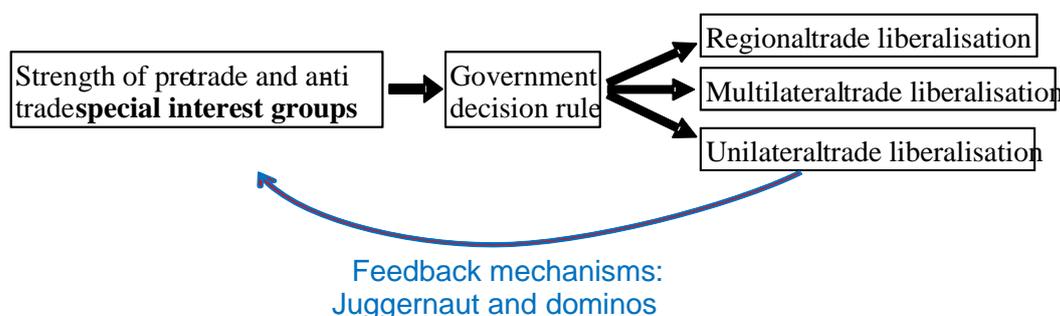
The next watershed year was 1973, which saw the launch of the Tokyo Round, as well as the massive broadening of European regionalism through the accession of the United Kingdom, Ireland and Denmark. This was also the year that the bilateral FTAs between the EEC and all the remaining members of the European Free Trade Association (EFTA) were signed.²⁴ By the end of 1973, all industrial tariffs on intra-Western European trade (almost half of total world trade at the time) had disappeared. All the nations that participated in these effective RTAs (the United States, Canada and Western European countries) were ardent multilateralists. The foremost distinction was not between regionalism and multilateralism, but between industrialised nations cutting tariffs and developing nations that did not.

This remarkable correlation of regional and multilateral tariff cutting calls for some explanation. After all, if we want to understand why the two are no longer marching together in the 21st century, we need to understand why they were in synch in the 20th century. I believe the answer lies in the so-called “juggernaut” and “domino” effects.²⁵

The schema in Figure 14 illustrates the basic reasoning. Tariff levels result from the political economy balance between anti-trade and pro-trade forces inside each nation (typically, import-competing firms are anti and exporting firms are pro). Tariff cutting occurs when something alters this balance. The key post-war novelty was the reciprocity principle contained in the GATT: the announcement that foreign tariffs would only be cut if domestic tariffs also were had a rallying effect. Domestic exporters added their voice to the tariff-cutting coalition, leading to a drop in each nation’s politically optimal tariff and causing the success of the GATT Rounds.

But this is not the end of the story.²⁶ As tariffs dropped reciprocally – either multilaterally or regionally – import-competing sectors shrank and wielded less influence on trade policy. Simultaneously, exporters in all nations expanded and wielded greater influence. As this tilted domestic tariff-setting balance toward more liberalisation in every nation, liberalisation tended to beget liberalisation. I dubbed this the “juggernaut” effect (Baldwin 1994, Chapter 2.5).

Figure 14. Juggernaut and domino effects



In the 20th century, multilateral reciprocal liberalisation occurs in phases over five to ten years (see above) and the necessary entry of exporters and exit of import competitors may take even longer. This naturally leads to tariff liberalisation that is episodic and synchronised. Once the entry and exit has occurred, governments find it

23. See Holmes (2005) and Foroutan (1998).

24. See Baldwin (1997).

25. Baldwin (1994) is the first presentation of the juggernaut effect; Baldwin (1993) is the first presentation of the domino effect. Bergstrand (1996) subsequently dubbed the first half of the domino effect “competitive liberalisation”.

26. See Baldwin (1994), Staiger (1995) and Baldwin and Robert-Nicoud (2008).

politically optimal to cut tariffs regionally and multilaterally. In short, once the tariff-cutting “juggernaut” gets rolling, the political economy momentum keeps it rolling until it crushes all tariffs in its path.²⁷

Another feedback mechanism is the “domino” effect, whereby the signing of one RTA tends to induce the signing of more RTAs. As FTAs reduce third-nation exports to the FTA nations, they stimulate third-nation exporters to engage in new political economy efforts to persuade their government to redress the new discrimination. In many cases, third-nation governments respond by signing new FTAs with one or both of the recently integrated partners.²⁸ The dominos continue to fall as each new RTA increases the trade diversion befalling excluded nations.

International political economy of 21st century trade liberalisation

Old-fashioned mercantilist exchanges of market access drove the political economy of 20th century RTAs. The international bargain was, “my market for yours”.²⁹ With 21st century RTAs, the basic international bargain is very different. It is “Northern factories for Southern reform”. This bargain is primarily bilateral, as the factories come from a specific nation, e.g. Japan, and the reform is done by a specific developing nation, e.g. Thailand. When a nation like the Philippines looks for Japanese factories, it goes to Tokyo to sign a deep RTA, not to Geneva to finish the Doha Round.

There is nothing new in this. Regionalism underpinned international supply chains in the 1960s and 1970s in North. US-centred production sharing was underpinned by the 1965 United States-Canada Auto Pact, while production sharing among West European nations was underpinned by the EEC Common Market. Japan, the only other manufacturing giant at the time, did very little production sharing and required no such disciplines.

This analysis helps explain why regionalism and multilateralism ceased to progress in tandem after the second unbundling. On the one hand, the juggernaut and domino effects ran out of tariff “fuel”. The GATT Rounds brought G7 nations’ tariffs down to very low levels and unilateral liberalisation did the same for developing nations’ tariffs. Moreover, most GATT members are not deeply involved in production sharing, so negotiating new disciplines in the WTO would be cumbersome and slow. On the other hand, negotiating deep RTAs was fairly straightforward, given the “factories for reform” nature of the political economy drivers.

In fact, the new development in the 21st century RTA saga has been mega-regionalism, as exemplified by the Trans-Pacific Partnership (TPP). If it concludes, it is likely to knit together many existing deep RTAs and extend their coverage to new bilateral relationships.

27. Juggernaut is a mispronunciation of the Hindu deity of the Puri shrine, Jagannath, whose chariot – an enormous and unwieldy construction – requires thousands to get rolling but was hard to stop once in motion. See Baldwin (1994, p.73) for the first presentation of the idea, Baldwin and Robert-Nicoud (2009) for formal modelling and Fugazza and Robert-Nicoud (2010) for empirical evidence.

28. See Baldwin (1993) for the original formulation of the domino theory, Baldwin (1997) for an early application and Baldwin and Jaimovich (2009) for a formal model. Egger and Larch (2008) and Baldwin and Jaimovich (2009) provide empirical support. On the theory, also see Bond and Syropoulos (1996), Freund (2000), Yi (1996), McLaren (2002), Levy (1997) and Krishna (1998).

29. As Cooper (1971, p. 410) puts it: “The principle of reciprocity is designed to hold out the promise of export gains to certain sectors of the economy, and thereby to establish a counterweight to those who will be hurt by increased imports. Reciprocity attempts to build pluralistic support for tariff reduction.” Well known to trade negotiators, this point was surely not novel to Cooper. Many made it subsequently, including Roessler (1978), Blackhurst (1979) and Baldwin (1980). For an early formal treatment, see Moser (1990) or Hillman, Long and Moser (1995). Grossman and Helpman (1995) brought the basic logic of these early papers to the attention of the broader community of trade academics.

6. Multilateralising deep regionalism: Practical issues

Since 21st century regionalism centres on quite distinct policies from the preferential tariffs that drove 20th century regionalism, the multilateralisation of 20th and 21st century regionalism should be quite different. To flesh out this observation, this section turns to practical concerns surrounding the multilateralisation of 21st century issues. To start with, we consider the nature of disciplines embodied in 21st century RTAs.

Multilateralising preferential agreements on services trade

Trade in services used to be called “trade in invisibles” – an outdated terminology reflecting the reason why services trade faces different barriers than goods trade. With rare exceptions, services trade is regulated behind the border, rather than at the border. See Borchert et al. (2012) for an inventory of services trade barriers.

In RTAs, services are classified by modes. Mode 1 comprises services which, like goods, are made in one nation and sold to another. Mode 2 is consumption abroad (e.g. tourism). Mode 3 is basically FDI, and Mode 4 is short-term migration. Twenty-first century RTAs frequently address services in modes 1 and 3, although Japanese deep RTAs usually include special provisions for temporary movement of key technicians and managers. For details on the rapid spread of services RTAs, see Miroudot et al. (2010), Fink and Jensen (2007) and Roy, Marchetti and Lim (2006).

Missing discrimination technology

A key feature in thinking about the multilateralisation of services trade is that discrimination across importers can be technically difficult.

Services are intangible products. Establishing where they were “made” is not an easy task (value added taxation of services faces the same challenge). Moreover, such origin-linked rules are frequently at odds with the announced regulatory goals. For example, prudential regulations for insurance providers should assure consumers that the insurance company is solid enough to pay claims when they arise. If the company is sufficiently solid, its nationality should not be an issue one way or the other.

Given these facts, nations typically regulate services trade by imposing requirements on service providers, rather than services themselves. For example, banks are only allowed to take deposits if they are subject to prudential regulations that meet national standards. The nation does not tax or limit deposit-taking, it simply forbids the bank from operating at all, or offering deposits.

Regulating services trade in this way creates huge loopholes in preferential access. For example, many provisions in the Japan-Thailand FTA apply to “enterprises of the other Party”, which it defines as: (i) a company constituted or otherwise organised under the law of the other Party and engaged in substantive business operations in that other Party; or (ii) in the case of the supply of a service through commercial presence, owned or controlled by natural persons of the other Party, or enterprises of the other Party in the sense of point (i). Thus, an American service provider that is active in Japan benefits from the Japan-Thailand bilateral agreement.

In a sense, the rules of origin for services in RTAs are “leaky”. For instance, while Sony can reasonably be considered a Japanese company, Sony USA is a US company. As Miroudot et al. (2010) put it, “rules of origin for services providers play an important role in minimising the distortions introduced by RTAs as firms from third-countries can benefit from the preferential treatment of RTAs through commercial presence in the territory of the parties.”

Due to these factors, Fink and Jansen (2009) argue that preferential agreements for services trade have not created a tangle of preferences, as happened with 20th century RTAs. In essence, most services FTAs have provided for “multilateralisation on autopilot”. The leaky rules of origin automatically multilateralise the preferential market access to a certain extent, as third-nation service providers can pay to establish a presence in one of the partner markets. Hence, unless all of the partners have equally stringent restrictions, the liberal rules of origin tend to lower all RTA partners’ MFN market access to that of the most liberal member (plus the extra establishment cost).

“Non-party MFN clauses”, a rarity in 20th century agreements, are yet another common feature of 21st century services RTAs. They allow the RTA parties to automatically enjoy any preferential treatment that either party extends to other nations. This measure, akin to multilateralisation on autopilot, avoids a tangle of bilateral privileges by automatically making the most liberal provision applicable to all parties.

Competition policy

Supply chain trade involves production abroad and this potentially exposes foreign firms to unfair competition. Trade agreements have long reflected this concern. The 1958 Treaty of Rome took this threat so seriously that it granted the European Commission direct regulatory powers on the matter. The worry was that anti-competitive practices (e.g. domestic firms colluding against imports) would offset the liberalising impact of lower tariffs.

For many years, competition policy featured mainly in European RTA provisions (Figure 8). Today, they are more prominent – especially with respect to state-owned firms, which often play a major role in emerging markets. Solano and Sennekamp (2006) reviewed the competition chapters of 86 (overwhelmingly North-South) FTAs, finding provisions on:

- adopting, maintaining, and applying competition laws
- co-ordination and co-operation between competition enforcement bodies
- addressing specific forms of anti-competitive behaviour
- competition principles reflecting core principles, including non-discrimination, due process and transparency
- excluding, or altering the recourse to, trade remedies
- dispute settlement
- special and differential treatment for developing countries.

Of the two main approaches to competition policy, the European Union’s is more prescriptive, focusing on persuading partners to embrace EU practices. The US approach centres on co-ordination and co-operation between the existing competition authorities, acting on their own regulations. A third way is exemplified in the Japan-Thailand FTA, which stipulates that each party shall promote fair and free competition by applying its own rules, without discriminating on the basis of nationality.

Very few RTAs have explicitly discriminatory provisions in the competition chapters, even if they do not specifically rule out preferential treatment. This is quite natural. Competition policy rests on domestic concerns over fairness or economic efficiency, so a firm’s nationality is irrelevant.

In short, the commitments included in RTAs have a multilateral effect, despite being codified in a bilateral agreement. For example, US firms in Turkey have the same rights before Turkish competition authorities as EU firms, even though Turkey’s competition policy resulted from bilateral agreements with the European Union (Kulaksizoglu, 2006). It is quite likely that this lack of discrimination is driven by the same issue that arose in services trade, i.e. that it is difficult to determine a corporate’s nationality. As a result, nations tend to adopt policies applicable to all corporate activity inside their borders, regardless of nationality. For this reason, RTA provisions on competition policy and state-owned enterprises are best viewed as unilateral pro-business policy reforms by developing nations. The RTA merely locks in an essentially autonomous non-discriminatory reform.

Investment: RTAs, BITs and capital movement provisions

The most visible aspect of GVCs is trade in parts and components. The second most obvious is FDI, whose protection is a core element of the package used by many developing nations to join international supply chains. FDI flows are protected by more than 2 800 bilateral investment agreements and 300 free-trade agreements with investment chapters (Berger, 2013). The rights provided include national treatment (in the WTO sense), fair and equitable treatment, and the freedom to move capital. Typically, these rights are enforceable before an international tribunal rather than national courts.

Two major approaches rule the global network of BITs (Berger, 2008). The European “admission model” (used by the main developed and developing nations, including China), protects investments only after the FDI meets the host country’s domestic laws and regulations. The North American “pre-establishment model” (used by the United States since the 1980s, Canada since the mid-1990s and Japan since 2000), impinges to a much greater extent on host nation prerogatives by restricting their screening powers. This restriction thus leads to greater ex ante openness toward FDI.

To feel comfortable setting up supply chain operations in the nation, investors need to feel confident they can control capital flows (e.g. capital expansion and contraction and profit repatriation). The global governance in this area is fragmented (Lupo Pasini, 2011). It involves an array of policies not commonly associated with trade (including capital controls and exchange restrictions). These are featured in the IMF rules, the WTO GATS rules (for service-related investments) and the OECD Code of Liberalisation of Capital Movements (for North-North flows).

To address this fragmentation and lock in liberal disciplines, many North-South deep RTAs include provisions on capital flows. The investment chapter of the Korea-United States FTA requires both parties to allow all in-and-out transfers related to the other party’s investment and explicitly prohibits currency market restrictions.

As with services and competition, investment assurances lack effective discrimination technology. To avoid relying on leaky rules of origin, the nationality of investments is not a concern in NAFTA. Its Article 1106 prohibits nations from imposing performance requirements (related to export, domestic content, etc.) on foreign investments – but the proscription applies to all foreign investors, not just those from other NAFTA nations.

Technical barriers to trade (TBTs)

Although TBTs are a serious issue, they are extremely difficult to liberalise in trade agreements. Baldwin (2000) points out that only two methods – hegemonic harmonisation and mutual recognition – have worked. The European Union and United States apply hegemonic harmonisation with their smaller trade partners. Mutual recognition is only possible among trade partners that profoundly trust each other’s regulatory systems (e.g. in the context of the European Union, EFTA and the Australia-New Zealand Closer Economic Relations Trade Agreement (ANZCER)).

Piermartini and Budetta (2007) surveyed a representative sample of 70 FTAs, 58 of which included TBT provisions. Beyond simple measures such as transparency and notification, they found three types of commitments: harmonisation of norms; mutual recognition of norms; and mutual recognition of certification procedures of each other’s norms.

Apart from the European and ANZCER arrangements, regional TBT agreements’ liberalising elements focused on mutual recognition of testing facilities in sectors such as pharmaceuticals and electrical equipment. In mutual recognition agreements (MRAs) signed between the European Union and the United States, for example, the European Union recognises the right of certain US laboratories to certify goods meeting EU norms (which can thus be sold in the European Union), while the United States recognises the right of certain EU laboratories to certify goods meeting US norms (which can thus be sold in the United States). This arrangement lowers costs, since all US-norm testing previously had to be performed in the United States and all EU-norm testing in the European Union.

As with all the other deep provisions, rules of origin make very little sense with regard to standards. One of governments’ basic motive for standards – to avoid all bad goods from reaching consumers – automatically makes them multilateral. If the goods meet the standard, their nationality is irrelevant as far as safety regulation is concerned. This lack of rules of origin effectively multilateralises TBT provisions in RTAs.

Take the MRA between the European Union and the United States. While the MRA appears to disadvantage Mexican firms in the EU market, the fact that Mexican firms can use US laboratories to prove their products’ conformity with EU norms considerably reduces the discriminatory effect. Compare the MRA with an FTA without rules of origin between nations A and B. Third nations – say, nation C – can still avail themselves of A’s duty-free access to B by trans-shipping goods via A. While C’s firms are at a relative

disadvantage compared to B's firms, it can never exceed the trans-shipment cost. Similarly, having a testing laboratory in the United States provides US firms with an edge over Canadian firms, but the advantage is limited to saving on Canadian firms' extra cost of having their products tested in the United States.

The case for more extensive multilateralisation of TBTs is subtle and indeed, worldwide TBT liberalisation may not be a good idea: valid arguments (e.g. judicial competition and differences in preferences and endowments) support allowing nations to set different regulations. There is no reason for South Africa and France to have the same standards, and therefore little reason for mutual recognition of norms. On the contrary, a common standard would likely harm both (Baldwin, 2000). Similarly, the wide gaps in income levels and governance capacity at the global level rule out liberalisation. European (or any other) norms might not be optimal for RoW, and vice versa (Bhagwati and Hudec, 1996).

Summary

A key lesson from the above discussion is that discrimination is difficult with regard to many 21st century trade disciplines – rules of origin either make no sense, or are “leaky” since nationality is difficult to pin down when it comes to services, companies and capital. This “leakiness” of the rules of origin dilutes the discriminatory effects of RTA provisions. The key upshot is that multilateralising 20th and 21st century regionalism present very different practical challenges.

7. Which measures should be multilateralised?

While governance at the global level is clearly most efficient where tariffs are concerned, this statement is less true for the deeper disciplines of 21st century regionalism. The question is, at what level should 21st century disciplines be set – globally, regionally, bilaterally, or unilaterally? In other words, what is the appropriate level of governance?

This question is not new. It arises within nations (e.g. should school curriculums be set at the national or sub-national level?) and across nations (e.g. should air passenger transport rules be set nationally, regionally or globally?).

The European Union has a long experience in this “allocation of competencies” – for instance, should anti-competitive behaviour be regulated at the European Union, national, or sub-national level? After struggling with the problem for six decades, it has largely adopted a set of principles to arrive at a set of answers.

The key EU principles are “subsidiarity” and “proportionality”. Subsidiarity means the European Union should only be in charge of those policies where EU-level action is more effective than at national, regional or local action. Proportionality means the European Union should be involved to the least extent necessary.

The EU answers are illuminating. Where external trade policy is concerned, the European Union harmonises policy to maintain the customs union. Likewise, it controls state aids policy to avoid beggar-thy-neighbour policies among members. Nations, however, are left entirely free to decide on corporate taxation.

“Fiscal federalism” as an analytic framework

Economists use the theory of fiscal federalism as an analytic framework to consider the governance-level problem. Rather than provide clear-cut answers, this theory helps approach the question intuitively and systematically.

Four key trade-offs permeate any consideration of the correct governance level.

Diversity of preferences: when people in different nations have very different preferences, a fully multilateralised, one-size-fits-all policy may not be optimal. For instance, nations differ radically in their ability and willingness to absorb foreign workers. A one-size-fits-all policy on temporary migration of workers will therefore most likely not improve the welfare of all WTO members.

Economies of scale: joint action by many nations can frequently reduce costs and/or increase effectiveness. The “Harmonized System” (HS) of trade classification is a good example. In principle (and in practice before the Second World War), each nation could create and use its own categorisation of imports. Such fragmentation

would, however, cause firms to incur extra costs as they attempt to justify their product's classification – e.g. to determine the applicable tariff – differently for each importing nation. This is a matter of scale economies. With a global coding system, firms can amortise the fixed cost of classifying their products over many more sales, thus lowering their average cost.

Since 1950, the World Customs Organization (WCO) has helped reduced fragmentation by maintaining the HS. In an interesting twist – which holds important lessons for multilateralising 21st century regionalism – harmonisation only goes so far. While the system identifies about 5 000 categories at the finest level of disaggregation (HS 6), nations are free to refine this further on a non-harmonised basis. The European Union, for example, classifies imports down to the eight-digit level, the United States to ten digits and Japan to nine digits. There is no harmonisation among these finer taxonomies.

Spillovers: many public policy choices involve positive or negative effects that cross sub-national and national boundaries. Take production subsidies: when a nation subsidises its firms' production, it generally imports less or exports more; this supply shift tends to lower prices, with a spillover effect onto firms in other nations. Thus, individually rational policies can lead to collective folly. This is a good argument for enacting disciplines at the global level, as happens with the WTO Agreement on Subsidies and Countervailing Measures. Yet spillovers can also be positive: innovation in one nation tends to spur growth in all, which is why the WTO explicitly allows R&D subsidies.

The very existence of spillovers does not mean a rule should be set at the global level. Lower-level governments can co-operatively take them into account, as with the voluntary setting of industrial standards. Global rule-setting may also not be ideal when preferences vary considerably. France and other nations insisted that the WTO treat "culture" differently than other commercial products. While the logic of mutual gains from freer trade in goods is clear even in cultural products (e.g. films), the diversity of preferences for maintaining national cultural production has led to the "cultural exception".

The trade-offs highlighted here assume that governments are well-intentioned. A very different line of thinking holds that policies should not be globally harmonised, so as to restrict governments' ability to exploit their own citizens.

Jurisdictional competition and democracy as a control mechanism: it is not uncommon for politicians and government officials to systematically favour politically powerful special-interest groups – e.g. by granting them tax breaks, subsidies and favourable laws – even to the detriment of average citizens. Hence, voters may be well advised to keep their governments in check through the two chief control mechanisms of jurisdictional competition and democracy.

Voters and firms can influence their government through either "voice" or "exit". Voice means operating within the system to change things. Exit – or jurisdictional competition – means to leave the jurisdiction imposing the policy, forcing decision makers to pay closer attention to public opinion. If a particular policy is globally harmonised, the exit option (which plainly applies more to firms than people) has no weight.

When policies are set at the national level, citizens and firms can directly influence policy choices, which is much more difficult when policies are set at the global level. In this sense, national policy setting provides them with more leverage to discipline their own governments' welfare-lowering actions.

In short, policy making at the national level tends to improve government by forcing nations to compete on policy choices and provide the "best value for money". Just as in the marketplace, competition can improve quality and reduce prices.

Current practices in light of the principles

While fiscal federalism provides a way of addressing the question of which 21st trade disciplines should be multilateralised, it does not provide clear answers. It is therefore instructive to study two long-standing WTO/GATT choices through the lens of this approach.

Since one nation's tariffs have direct spillover effects on market conditions in other nations, tariff disciplines are best set at the global level to reduce negative spillovers – which is why the first "T" in "GATT" stands for tariffs. This classic solution, however, does not fit all cases. GATT did not insist every nation have

the same tariffs, but rather that each nation's tariff-setting mechanism respected certain rules, the foremost being non-discrimination (Article 1). The rules on the deviations contained in RTAs are set at the global level (Article 24). Heterogeneous preferences and conditions play a role, in that the rules for discrimination are laxer for PTAs among developing nations than for PTAs involving at least one developed nation.

By contrast, the GATT general principles governing non-discrimination behind the border explicitly excluded government procurement (Article 3.8, "National treatment"). While the economic benefits of non-discrimination apply as directly to government purchases as they do to private purchases, national preferences varied too widely to justify a single discipline at the global level. Instead, a subset of WTO members banded together to adopt the Agreement on Government Procurement, whose one-size-fits-all disciplines apply only to its signatories.

The lack of spillovers

The spillovers are much less obvious when it comes to most deep RTA provisions. As seen in Section 0, there is little empirical evidence for trade diversion. The practical considerations discussed in Section 4 help explain why these PTAs are not, in practice, very preferential. Many deep RTA provisions come without obvious rules of origin, or with leaky ones.

The deep RTA provisions involving property rights protection are similarly leaky. If an FTA between two nations provides investment protection to firms operating in each other's markets, third-nation firms may gain the same protection by registering – or taking on a partner – in one of the RTA nations. Moreover, the RTA signers may change their national laws to protect the property rights of all investors, with the fundamental policy goal of fostering inward investment by all advanced technology firms, not just those from the partner nation.

In other cases, the deep RTA provisions establish no preferences at all. For example, the IPR chapter in the United States-Korea RTA requires the parties to accede to various existing treaties.³⁰ Since these treaties are open, the FTA is not creating a negative trade diversion spillover; it is merely a vehicle for locking in domestic reforms.

Network externalities and "standards competition" analysis

The preceding analysis suggests there is a need to rethink approaches to multilateralising deep RTA provisions. One way is to consider 21st century rules in light of network externalities and standards competition, an economic reasoning that permeates the analysis of high-technology industries, such as apps for smart phones. Apps are marked by clear network externalities. If lots of people use Apple phones, developers will write lots of iPhone apps, and the availability of so many apps makes Apple phones more attractive. The same is true for Windows-based smartphones and related apps.

Upon reflection, it is clear that this perspective is useful when thinking about the costs and benefits of multilateralising 21st century regionalism. That is, the gains in multilateralising deep RTA provisions stem not from eliminating negative spillovers, but rather from realising greater network externalities. For example, we know there are about 2 800 BITs in existence, divided into two basic categories – the European and US models. Here, the notion of network externality relates to the idea that home and host nations would be better off with a single rulebook. In the case of smart phone apps, the existence of multiple operating systems makes it more difficult to do business with friends using different systems.

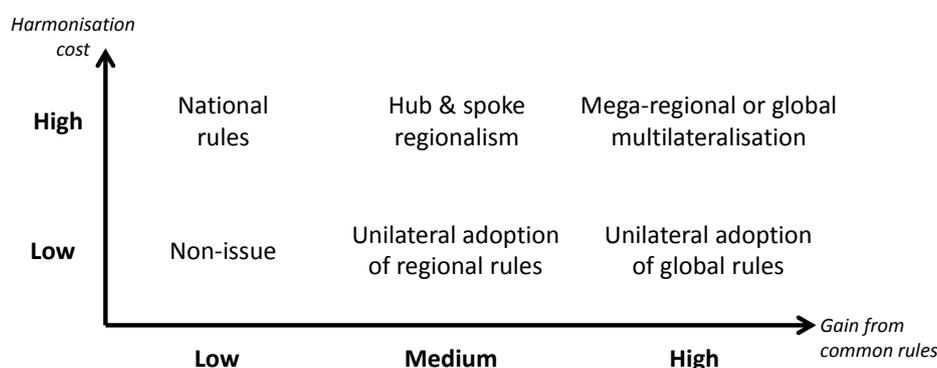
30. The treaties listed are the Paris Convention for the Protection of Industrial Property, Berne Convention for the Protection of Literary and Artistic Works, Convention Relating to the Distribution of Programme-Carrying Signals Transmitted by Satellite, Protocol Relating to the Madrid Agreement Concerning the International Registration of Marks, Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure, International Convention for the Protection of New Varieties of Plants, Trademark Law Treaty, World Intellectual Property Organization Copyright Treaty and Performances and Phonograms Treaty.

Spillovers, harmonisation costs and level of governance

While network externalities tend to favour multilateralisation at the global level, there is a trade-off. Multiple rulebooks exist on items such as IPR, investment assurances, BITs and capital mobility strictures. Transforming these into common rules entails “harmonisation costs”. The question of multilateralisation must balance these costs with the benefit of greater network externalities.

Figure 15 helps organise the analysis. The diagram plots the cost of harmonising rulebooks on the vertical axis; it illustrates the benefit of harmonisation – namely, the importance of network externalities – on the horizontal axis. In principle, while the costs and benefits could (given perfect data) be quantified in dollars, we merely illustrate them here schematically as “high” or “low”. An important dimension, which does not fit easily into the diagram, is the localness of the network externalities. Take BITs: since most trade and investment – especially supply chain trade – is regional (Gamberoni et al., 2010), agreeing on common rules affecting supply chain trade and/or the associated investment will likely have regional rather than global spillovers. More precisely, the strength of the spillovers from harmonisation will diminish over distance. In the absence of an explicit “distance” axis, we use the “high but most regional” category to indicate mostly regional or local spillovers.

Figure 15. Schema for thinking about multilateralisation levels



Source: Author's elaboration.

If we start at the bottom row (left side) with the case of low benefits of multilateralisation (i.e. low gains from common rules) and low costs of harmonisation, the sensible option is to ignore the issue. Each nation does what it wants and the cost of the missed opportunity of setting common rules at the bilateral, regional, or global level is minimal. In the case of locally provided services, e.g. hotel services, each nation has its own regulations and multinational hotel chains adapt.

Moving to the right, the next case entails medium-sized gains to common rules and low costs. The network externalities are regional, rather than global, in reach. For example, the NAFTA approach to BITs started as the US approach; while Canada and Japan unilaterally adopted it, Europeans did not. At the same time, the economic dominance of the European Union has led its trade partners to tend toward the European model in their own BITs.

The right-most case in the bottom row involves low costs and high benefits to global harmonisation of rules. While examples of policies in this category abound, they are uncontroversial and therefore rarely discussed. Most WCO and much International Organization for Standardization work involves identifying “focal point” equilibriums for global co-operation and helping nations co-ordinate on them. The USB connector is a good example. Adoption of a common standard provided a huge benefit to global trade, with low harmonisation costs.

Turning to the top row, the issues get more contentious. The high harmonisation costs may stem from the diversity of preferences across nations (e.g. income tax rates), or from more mechanical issues arising in

complex regulatory regimes (e.g. banking regulation, where the Basel Committee is finding it massively complex to move toward global standards).

The top-left category, “national rules,” contains most national economic policy issues, where spillovers are small and costs of harmonisation are high. Nations set their own rules and no one complains, since there are few spillovers. In the middle-top category, “hub and spoke regionalism”, the idea is that while the costs are high, harmonisation offers major positive network externalities. In such cases, nations generally require some carrot-and-stick mechanism – such as the domino effect from bilateral RTAs – to perform regional harmonisation.

A good example here is the spread of the NAFTA rules of origin. In Asia and Europe, rules of origin tend to be based on value added principles. Firms – which already have to gather most of the required data for value added tax purposes – find them clear, easy to implement and efficient. The United States, however, has no value added tax and has adopted the “change of chapter headings” approach to rules of origin. Australia previously used the value added approach, but switched to the US system when it signed an FTA with the United States. It then convinced New Zealand, its long-standing FTA partner, to switch to the US system to avoid duplicative administration. This regional harmonisation happened via hub and spoke – the carrot is tariff preferences and extra investment from the hub (in this example, the United States), while the stick is the fear of being harmed by trade and investment diversion. The hub has very little interest in turning this process into a plurilateral discussion and the spokes have little commercial interest in discussing the issue with the other spokes orbiting around the hub’s gravitational pull.³¹

The top-right entry in the schema represents the hardest case – high cost and high benefits – which requires serious inducements to achieve harmonisation. The Agreement on Government Procurement (GPA) of the WTO is a good example. The gains from freer competition for government purchases are large, both in terms of cost savings and mercantilist export. However, the resistance from sheltered providers is equally fierce and requires serious leverage to overcome.

The method currently applied is mega-regionalism. The United States in particular is orchestrating plurilateral talks in the Asia-Pacific (through the TPP) and Atlantic (through the Trans-Atlantic Trade and Investment Partnership (TTIP)) regions. Other talks are underway between Japan and the European Union, Canada and the European Union and Canada and Japan. If these succeed, multilateral rules would most likely resemble TPP and TTIP rules and global multilateralisation of deep RTA rules would essentially entail nations adopting the TPP-TTIP rulebook unilaterally. If the mega-regionals fail, global level negotiations could eventually multilateralise deep RTA provision.

A tentative list of deep RTA provisions to be multilateralised

Based on this reasoning, I propose that a shortlist of deep RTA provisions should include deeper commitments on TRIMs, TRIPs, customs and GATS, as well as new commitments on competition policy, IPR, investment, movement of capital and the approximation of legislation. But which of these provisions should be multilateralised at the global level?

Investment rules: As argued in the introduction, today’s world of GVCs blends trade in goods, services, investment and IP. As the WTO sets the basic rules on trade in goods and services, it seems natural it should also set the basic rules on investment – but not in the form of a one-size-fits-all policy. The first step would be to agree on basic principles, such as national treatment and third-party arbitration. WTO members would then harmonise specific wording according to regional practices. The inspiration here is the GATS text, which sets out some general guidelines, but leaves future multilateral trade and regional agreements to fill in most of the fine print.

The gains from multilateralising investment rules would include bountiful network externalities from reducing the tangle created by thousands of BITs. Moreover, thanks to the close analogy between trade

31. A similar pattern occurred in the Central American Common Market (CACM), when the United States signed bilaterals with the Dominican Republic and CACM members using value-added rules among themselves. After signing with the United States, CACM nations agreed to recognise origin based on either the US chapter-heading rules or the CACM value-added rules.

diversion and investment diversion (Baldwin, Forslid and Haaland, 1996), the idea that multilateralisation would reduce negative spillovers applies as much to investment as it does to tariffs.

Another argument in favour of global investment governance is its impact on perceived systemic fairness. As Lawrence (1996) argued, “Major regional arrangements could be dominated by considerations of market power rather than the principles of a liberal trading order.” Thus, nations other than the largest global players would have to either sign up and play by their rules or be left out in the cold. A multilateral agreement on investment provisions could yield a more equitable outcome. The GATT (Article 24) restricts power asymmetries in RTA talks. Both parties must arrive at zero tariffs on substantially all trade, which limits the bigger nation’s ability to force an asymmetric deal on its smaller partner. Similarly, a global deal on investment provisions could strengthen the hand of small developing nations faced with large FDI emitters.

Finally, preferences have converged massively since globalisation’s second unbundling. For years, quasi-Marxist thinking led developing nations to treat FDI as an attempt to steal the family silver. Now, they view FDI as the most promising way to jump on the supply chain industrialisation escalator. These changed circumstances have led the United Nations Conference on Trade and Development, the OECD, the International Chamber of Commerce and Asia-Pacific Economic Cooperation to take steps toward multilateralising investment disciplines.

Customs co-operation: The obvious network externalities arising from harmonised customs procedures become especially important for 21st century trade, where delays of days (or even hours) can create cascading problems for firms along the supply chain. Discussions on trade facilitation – the only area where the WTO is making some progress – have recognised this point.

Other issues: The diversity of preferences among supply chain participants would seem to limit the scope for one-size-fits-all rules in other disciplines. In some disciplines, like competition policy, the disagreement is mainly between the United States and the European Union. In others, like IPR, the divide is mainly between the giant emerging nations (e.g. China and India) who have the market power to attract offshored industrial jobs without offering strong IPR protection and all other nations. Nations with advanced technology favour strong IPR, while small developing nations are willing to embrace the disciplines in exchange for Northern factories.

Multi-tier multilateralisation

Given the extreme difficulties of arriving at global one-size-fits-all disciplines, Baldwin and Thornton (2008) suggest multi-tier multilateralisation.

The first – and loosest – tier would be for the WTO to establish voluntary best-practice guidelines for new RTAs and modifying existing RTAs. The idea is to encourage nations to consider the impact of their agreements on non-party WTO members and provide model language to reduce unnecessary wording differences across RTAs. This modest step could involve a hierarchy of best-practice guidelines for North-North, North-South and South-South RTAs. The inspiration here is Article 24 of the WTO, which imposes stricter standards on RTAs between developed nations than the Enabling Clause does on RTAs between developing nations. The WTO could also consider a “super-developed” nation standard – a sort of gold standard of 21st century RTAs.

The second tier would consist in setting a set of minimum principles, e.g. national treatment, third-party MFN and transparency. The idea is to take advantage of the reduced diversity of preferences among WTO members, many of which have signed RTAs with commonalities (albeit different details). This is akin to nations agreeing to harmonise their import classifications at the HS-6 level, but having leeway to specify more detailed categories. The first step here would be to launch a major legal research project documenting the similarities or differences among the deep RTA provisions in practice.

8. Conclusions

Trade and trade agreements used to be relatively simple.

- Trade primarily meant trade in “made-here-sold-there” goods.
- Twentieth-century regional and multilateral trade agreements dealt primarily with tariffs and other border-barriers.
- The key purpose of trade and trade agreements for governments was to help their firms sell things.

The internationalisation of production networks from high-wage to low-wage nations – call it the GVC revolution – changed all this. The trade system is being used to *make* goods, rather than simply *sell* them. The resulting transformation of international commerce and commercial policy has triggered a paradigm shift.

First, the definition of trade has been stretched. International commerce involves richer, more complex and more interconnected exchanges. Put simply, the goods, services, ideas, people, know-how and capital that used to move only within rich-nation factories are now crossing borders. This turned GVC-linked trade into a nexus of trade, services, investment and IP.

Second, more complex commerce required more complex trade agreements. Deals that promote trade in “made-everywhere-sold-there” goods must address disciplines related to the *making* as well as the *selling* of goods. Trade agreements must provide assurances that (i) GVC-linked flows of goods, services, and capital can easily cross borders; and (ii) GVC-linked tangible and intangible property rights are respected.

Third, governments’ motives have shifted. For rich nations, 21st century agreements underpin international supply chains that are critical to their firms’ competitiveness. The motives for most developing nations are quite different. As GVC participation is the 21st century’s fast lane to industrial development, 21st century trade agreements are keystones in many nations’ development strategies. Oversimplifying to make the point, the bargain behind 21st century trade agreements is Northern factories for Southern reform. For 20th century agreements, the bargain was “my-market-for-yours”.

Fourth, the GVC revolution transformed trade governance. The complex and fundamentally bilateral nature of the factory-for-reform bargain swung the centre of gravity decisively away from multilateral governance. Ad hoc governance structures have materialised to undergird international production networks. Since GVC networks are mostly regional, the governance responses have mostly been regional. The main elements are BITs, deep RTAs between advance technology “offshorers” and low-wage “offshorees”, as well as massive unilateral policy reform by developing nations.

The rise of mega-regional arrangements threatens to cement this new structure into place and further erode WTO centrality. While the WTO remains relevant to 20th century trade, global rules for 21st century trade are being written in the TPP, TTIP, TISA and the like.

The real threat to multilateralism

The rise of 21st century regionalism has been good for world trade. Despite slow progress by the WTO, trade and trade opening have boomed. Thinking ahead, however, it is clear that global trade governance faces a historical turning point. The current trajectory seems certain to undermine the WTO’s centrality, with the mega-regionals taking over as the main loci of global trade governance. Without reforms that bring existing deep RTA disciplines under the WTO’s aegis and facilitate development of new disciplines inside the WTO, the trend will continue. At best, the WTO will continue to thrive as the institution underpinning 20th century trade flows.

This is not the only scenario. WTO centrality could erode beyond the tipping point where nations ignore WTO rules, since everyone else does. The inability of the WTO to update its rules undermines the authority of the dispute settlement mechanism. Judges are increasingly forced to make rulings based on previous judges’ rulings, instead of on agreements negotiated by consensus. Danger lies down this road. A WTO that cannot finish negotiations and cannot effectively adjudicate would be moribund. The GATT/WTO would go down in history as a 70-year experiment where world trade was rules-based instead of power-based.

This darker scenario runs the risk of throwing global trade governance back toward a 19th century Great Powers arrangement. Back then, dispute settlements and trade agreements arose from reciprocal negotiation when two Great Powers were involved, but from exercises in pure economic muscle when Great-Power nations dealt with other nations. This should worry all world leaders. In the first half of the 19th century, attempts by incumbent Great Powers to impose rules on emerging powers paved the road to humanity's greatest follies – World War-I and World War-II.

Multilateralising 21st century regionalism: A better way forward

One programme of preventive action would be to work toward multilateralising the deeper disciplines in today's deep RTAs and BITs. Many of these are embodied in national laws. Many are endowed with a public-good nature, in the sense that they facilitate trade with all nations, not just the members of the RTA which brought about the reform. While questions of consistency seem secondary for such measures, for other measures – e.g. IP protection or investors' rights – the various deeper disciplines do not appear clearly compatible.

Distinguishing the various categories of disciplines is an important task for trade scholars and governments. The centrality of the WTO is not in peril if the various deep RTAs turn out to have implemented reforms that are consistent with each other. Such disciplines might easily be multilateralised with WTO agreements (e.g. the GATS) or plurilateral agreements (e.g. the Government Procurement Agreement). The disciplines that are creating mutually inconsistent rules are more of a problem and need to be identified.

Part of this exercise will be to identify which deeper disciplines are more efficiently organised at the global level and which are best set at the regional or national level. As discussed above, economic theory on the allocation of tasks to various levels of government (fiscal federalism) could be used to determine which of the deeper measures belong in the WTO and which are more appropriately dealt with in RTAs and/or national legislation. Again, this is an open question for trade scholars, governments and practitioners.

More modest versions of 21st century multilateralisation can also be envisioned. The WTO could develop some basic guidelines for deeper provisions in RTAs, akin to those on tariffs and services in the GATT and GATS. For example, the GATS provides a few basic guidelines for services FTAs, but the FTAs fill in the details for market opening. Even such very basic guidelines are completely absent when it comes to GVC-linked provisions, such as competition policy, rights of establishment, FDI-linked capital flows, IPR and the like.

The way forward

To be practical, I suggest that the WTO start multilateralising 21st century regionalism by addressing investment rules and customs co-operation. Any further action will face enormous challenges, as WTO members find it almost impossible to agree on one-size-fits-all disciplines.

As an alternative, the WTO should consider multi-tier multilateralisation (Baldwin and Thornton, 2008). The first tier – establishing voluntary best-practice guidelines for new RTAs – would encourage nations to consider the impact of their agreements on non-party WTO members and help reduce differences in wording (and thus interpretations) across RTAs. The hierarchy of best-practice guidelines – tailored to North-North, North-South and South-South RTAs – would allow for developmental differences.

The second tier would involve agreeing on basic principles – including national treatment, third-party MFN and transparency – already widely included in deep RTAs.

Whatever happens, the global system of trade governance will be transformed by the end of this decade. As in the 1930s and 1940s, global trade rules are being written by a handful of powerful nations. On the current trajectory, mega-regionalism will sideline multilateralism and undermine the centrality of the WTO. Multilateralising these new rules would help ensure that today's situation ends up like the latter half of the 20th century instead of the first half.

References

- Acharya, R., et al. (2011), "Landscape", in J.P. Chauffour and J.C. Maur (eds.), *Preferential Trade Agreement Policies for Development: A Handbook*, The World Bank, Washington DC.
- Amador, J. and S. Cabral (2009), "Vertical specialization across the world: A relative measure", *The North American Journal of Economics and Finance*, Vol. 20/3, pp. 267-280
- Ando, M. and F. Kimura (2005), "The formation of international production and distribution networks in East Asia", in T. Ito and A. Rose (eds.), *International trade in East Asia, NBER-East Asia seminar on economics*, Vol. 14, The University of Chicago Press, Chicago. First version, *NBER Working Paper* 10167.
- Antras, P. and F. Foley (2011), "Regional Trade Integration and Multinational Firm Strategies", in R. Barro and J.-W. Lee (eds.), *Costs and Benefits of Economic Integration in Asia*, Cambridge University Press, Cambridge, UK.
- Athukorala, P.-c. and N. Yamashita (2006), "Production fragmentation and trade integration: East Asia in a global context", *The North American Journal of Economics and Finance*, Vol. 17/3, pp. 233-256.
- Bagwell, K. and R. Staiger (1999), "An Economic Theory of GATT", *American Economic Review*, Vol. 89/1, pp. 215-248.
- Baldwin, Richard (2013), "Trade and industrialisation after globalisation's second unbundling: How building and joining a supply chain are different and why it matters", in R. Feenstra and A. Taylor (eds.), *Globalization in an Age of Crisis: Multilateral Economic Cooperation in the Twenty-First Century*, University of Chicago Press, Chicago.
- Baldwin, Richard (2012a), "Trade and Industrialisation after Globalisation's Second Unbundling: How Building and Joining a Supply Chain are Different and Why it Matters," in *Globalization in an Age of Crisis: Multilateral Economic Cooperation in the Twenty-First Century*, R. Feenstra and A. Taylor (eds.), University of Chicago Press, Chicago.
- Baldwin, Richard (2012b), "WTO 2.0: Global governance of supply-chain trade", *CEPR Policy Insight*, No. 64.
- Baldwin, Richard (2011a), "21st Century Regionalism: Filling the gap between 21st century trade and 20th century trade rules", *CEPR Policy Insight*, No. 56.
- Baldwin, Richard (2011b), "Trade and industrialisation after globalisation's 2nd unbundling: How building and joining a supply chain are different and why it matters", *NBER Working Paper* 17716.
- Baldwin, Richard (2010), "Unilateral tariff liberalisation", in *The International Economy, Journal of The Japan Society of International Economics*, No.14, pp. 10-43. Also *NBER Working Paper* 16600, 2010.
- Baldwin, Richard (2008), "Managing the noodle bowl: The fragility of East Asian regionalism", *The Singapore Economic Review*, Vol. 53/03, pp. 449-478.
- Baldwin, Richard (2006a), "Multilateralising Regionalism: Spaghetti Bowls as Building Blocs on the Path to Global Free Trade," *The World Economy*, Vol. 29/11, pp. 1451-1518.
- Baldwin, Richard (2006b), "Managing The Noodle Bowl: The Fragility Of East Asian Regionalism," *CEPR Discussion Papers* 5561, also published in *Singapore Economic Review*, Vol. 53/3, pp. 449-478, 2008.
- Baldwin, Richard (2006c), "Globalisation: The great unbundling(s)", in *Globalisation challenges for Europe*, Secretariat of the Economic Council, Finnish Prime Minister's Office, Helsinki, pp 5-47, <http://www.vnk.fi/hankeet/talousneuvosto/tyo-kokoukset/globalisaatioselvitys-9-2006/en.jsp>.
- Baldwin, Richard (2006d), "The euro's trade effects", *ECB Working Paper Series*, No. 594, March.
- Baldwin, Richard (2000), "Regulatory Protectionism, Developing Nations, and a Two-Tier World Trade System", *Brookings Trade Forum*, 2000.
- Baldwin, Richard (1997), "The Causes of Regionalism", *The World Economy*, Vol. 20/7, pp. 865-888. Also CEPR (Centre for Economic Policy Research) Discussion Paper 1599.

- Baldwin, Richard (1996), "Investment Creation and Diversion in Europe", *The World Economy*, Vol. 19/6, pp. 635-659.
- Baldwin, Richard (1993), "A Domino Theory of Regionalism", *NBER Working Paper No. 4465*, September; eventually published as Chapter 2 (pp. 25-48) in Baldwin, Haaparanta and Kiander (eds.), *Expanding Membership in the European Union*, Cambridge University Press, Cambridge, UK.
- Baldwin, Richard and Harry Flam (1997), "Enlargement of the European Union: The Economic Consequences for the Scandinavian Countries", *CEPR Occasional Paper OP16*.
- Baldwin, Richard and Patrick Low, eds. (2009), *Multilateralizing Regionalism: Challenges for the Global Trade System*, Cambridge University Press, Cambridge, UK.
- Baldwin, Richard and Philip Thornton (2008), "Multilateralising Regionalism", CEPR Policy Report.
- Baldwin, Richard and Philippe Martin (1999), "Two waves of globalisation: superficial similarities, fundamental differences", *NBER Working Paper 6904*. Published as Chapter 1, pp. 3-59, in H. Siebert (ed.), *Globalisation and Labour*, J.C.B. Mohr, Tübingen, 1999.
- Baldwin, Richard, Simon Evenett and Patrick Low (2009), "Beyond tariffs: multilateralizing non-tariff RTA commitments", in Richard Baldwin and Patrick Low (eds.), *Multilateralizing Regionalism: Challenges for the Global Trade System*, Cambridge University Press, Cambridge, UK.
- Baldwin, Robert (1980), "The Economics of the GATT", *Issues in international economics*, Vol. V, Oriel Press, London.
- Batra, R. N. and F.R. Casas (1973), "Intermediate Products and the Pure Theory of International Trade: A Neo-Heckscher-Ohlin Framework", *American Economic Review*, Vol. 63/3, pp. 297-311.
- Bems, R., R. Johnson and K. Yi (2010), "Demand spillovers and the collapse of trade in the global recession", *IMF Economic Review*, Vol. 58/2, pp. 295-326.
- Berger, A. (2008), "China and the Global Governance of Foreign Direct Investment – The Emerging Liberal Bilateral Investment Treaty Approach", *DIE Discussion Paper*, Vol. 10/2008, German Development Institute.
- Berger, A. (2013), "Do we really need a multilateral investment agreement?", *DIE Briefing Paper*, Vol. 9/2013, German Development Institute.
- Berger, A., et al. (2011), "More stringent BITs, less ambiguous effects on FDI? Not a bit!", *Economics Letters*, Vol. 112, pp. 270-272.
- Berger, A., et al. (2010), "Do trade and investment agreements lead to more FDI? Accounting for key provisions inside the black box", *International Economics and Economic Policy*, Vol. 7/1.
- Bhagwati, J. (2008), *Termites in the Trading System: How Preferential Agreements Undermine Free Trade*, Oxford University Press, Oxford.
- Blackhurst, R. (1979), "Reciprocity in trade negotiations under flexible exchange rates: Trade and payments adjustments under flexible exchange rates", in J. Martin and A. Smith (eds.), *Trade and payments adjustments under flexible exchange rates*, Macmillan, London.
- Blinder, A. (2006), "Offshoring: The next industrial revolution?", *Foreign Affairs*, March/April 2006.
- Borchert, I., B. Gootiiz and A. Mattoo (2012), "Policy barriers to international trade in services: evidence from a new database", *Policy Research Working Paper Series 6109*, The World Bank, Washington DC.
- Brühlhart, M. (2009), "An Account of Global Intra-Industry Trade, 1962-2006", *World Economy*, Vol. 32/3, pp. 401-459.
- Coates, D. and R. Ludema (2001), "A theory of trade policy leadership", *Journal of Development Economics*, Vol. 65/1, pp. 1-29.
- Conconi, P. and C. Perroni (2009), "Do credible domestic institutions promote credible international agreements?", *Journal of International Economics*, Vol. 79(1), pp. 160-170.
- Cooper, Richard (1971), "Third World Tariff Tangle", *Foreign Policy*, Vol. 4, pp. 35-50.

- Cañas, J. and R. Coronado (2002), “Maquiladora Industry: Past, Present and Future”, *Business Frontiers*, Issue 2, Federal Reserve Bank of Dallas.
- Damuri, Y. (2013), “Production Sharing Practice and the 21st Century Regionalism”, Chapter 2 in PhD Thesis, Graduate Institute, Geneva. Also *CTEI Working Papers*, CTEI-2012-4.
- Daudin, G, C. Riffart and D. Schweisguth (2011), “Who Produces for Whom in the World Economy?”, *Canadian Journal of Economics*, Vol. 44/4, pp. 403-1437.
- Deardorff, A. (2005), “Ricardian comparative advantage with intermediate inputs,” *North American Journal of Economics and Finance*, Elsevier, Vol. 16/1, pp. 11-34.
- DFAIT (2012), “Canada-Japan Free Trade Agreement Negotiations”, Department of Foreign Affairs and International Trade Canada, <http://www.international.gc.ca>.
- Duer, A. et al. (2012), “The Design of International Trade Agreements: Introducing a New Database”, paper prepared for the 25th conference of the German Political Association, 24-28 September 2012.
- Economist (2012), “A third industrial revolution”, *Economist Magazine Special Report*, 21 April 2012.
- Egger, P. and V. Merloz (2012), “BITs Bite: An Anatomy of the Impact of Bilateral Investment Treaties on Multinational Firms”, *Scandinavian Journal of Economics*, Vol. 114/4, pp. 1240-1266.
- Egger, P. et al. (2011), “The Trade Effects of Endogenous Preferential Trade Agreements”, *American Economic Journal: Economic Policy*, Vol. 3/3, pp. 113-43.
- Ethier, W.J. (1982), “National and international returns to scale in the modern theory of international trade”, *American Economic Review*, Vol. 72/3, pp. 389-405.
- EU-US High Level Working Group on Jobs and Growth (2012), “Interim Report to Leaders from the Co-Chairs EU-US High Level Working Group on Jobs and Growth”, <http://trade.ec.europa.eu>, Tradoc_149557.pdf.
- Feenstra, R.C. and G.H. Hanson (1996), “Globalization, Outsourcing, and Wage Inequality”, *American Economic Review* 86, pp. 240-45.
- Feenstra, R. and A. Taylor, eds. (2013), *Globalization in an Age of Crisis: Multilateral Economic Cooperation in the Twenty-First Century*, University of Chicago Press, Chicago.
- Francois, J. (1990), “Trade in Nontradables: Proximity Requirements and the Pattern of Trade in Services”, *Journal of Economic Integration*, Vol. 5/1, pp. 31-46.
- Freund, C. (2010), “Third-Country Effects of Regional Trade Agreements”, *World Economy*, Vol. 33/11, pp. 1589-1605.
- Fugazza, M. and N. Alessandro (2011), “Measuring preferential market access”, *MPRA Paper 38565*, Munich.
- Fujita, M., P. Krugman and A. Venables (1999), *The Spatial Economy: Cities, Regions and International Trade*, MIT Press, Cambridge (Mass.).
- Fukao, K., H. Ishito and K. Ito (2003), “Vertical intra-industry trade and foreign direct investment in East Asia”, *Journal of the Japanese and International Economies*, Vol. 17/4, pp. 468-506.
- Gamberoni, E., R. Lanz and R. Piermartini (2010), “Timeliness and contract enforceability in intermediate goods trade”, *Policy Research Working Paper Series*, Vol. 5482, The World Bank, Washington.
- Gereffi, G. (2001), “The Value of Value Chains: Spreading the Gains from Globalisation”, *IDS Bulletin*, Vol. 32/3.
- Grossman, G. and E. Rossi-Hansberg (2008), “Trading Tasks: A Simple Theory of Offshoring”, *American Economic Review*, Pittsburgh.
- Grossman, G. and E. Helpman (1995), “Technology and trade”, in *Handbook of International Economics*, Vol. 3, North-Holland, Amsterdam, pp. 1279-1337.
- Grubel, H.G. and P.J. Lloyd (1975), *Intra-industry trade: the theory and measurement of international trade in differentiated products*, Wiley, New York.

- Hanson, G. (2012), "The Rise of Middle Kingdoms: Emerging Economies in Global Trade", *NBER Working Paper*, No. 17961.
- Hillman, A., N. Van Long, and P. Moser (2008), "Modelling Reciprocal Trade Liberalization: The Political-economy and National-welfare Perspectives", *Swiss Journal of Economics and Statistics (SJES)*, Vol. 131/III, pp. 503-515.
- Hoekman, B. (1993), "Multilateral trade negotiations and coordination of commercial policies", in R. Stern (ed.), *The multilateral trading system: Analysis and options for change*, Harvester-Wheatscheaf, New York.
- Horn, H., P.C. Mavroidis and A. Sapir (2009), "Beyond the WTO? An Anatomy of EU and US Preferential Trade Agreements", *CEPR Discussion Papers 7317*, C.E.P.R. Discussion Papers.
- Hufbauer, G., T. Moran and L. Oldenski (2013), *Outward Foreign Direct Investment and US Exports, Jobs, and R&D: Implications for US Policy*, Peterson Institute of International Studies, Washington DC.
- Hummels, D., J. Ishii and K.-M. Yi (2001), "The nature and growth of vertical specialization in world trade", *Journal of International Economics*, Vol. 54/1, pp. 75-96.
- Hummels, D., J. Ishii and K.-M. Yi (1999) "The nature and growth of vertical specialization in world trade," *Staff Reports 72*, Federal Reserve Bank of New York, New York.
- Hummels, D., D. Rapoport and K.-M. Yi (1998), "Vertical specialization and the changing nature of world trade", *Economic Policy Review*, Vol. 4/2, Federal Reserve Bank of New York, New York, pp. 79-99.
- IDE-JETRO (2006), "How to make the Asian input-output Tables", Institute of Developing Economies, March 2006.
- Inomata, Satoshi (1995), "Asian International Input-Output Table – the significance of compilation", *Ajia-keizai*, Vol. 36/8, IDE Tokyo.
- Johnson, R. C. and G. Noguera (2012c), "Fragmentation and trade in value added over four decades", online pdf, January.
- Johnson, R. C. and G. Noguera (2012a), "Proximity and Production Fragmentation," *American Economic Review*, American Economic Association, Vol. 102/3, pp. 407-11, May 2012.
- Johnson, R. C. and G. Noguera (2012b), "Accounting for intermediates: Production sharing and trade in value added," *Journal of International Economics*, Elsevier, Vol. 86/2, pp. 224-236.
- Jones, R.W. and H. Kierzkowski (1990), "The Role of Services in Production and International Trade: A Theoretical Framework", in R. Jones and A. Krueger (eds.), *The Political Economy of International Trade*, Basil Blackwell, Oxford.
- Jones, R. (1980), "Comparative and Absolute Advantage", *Swiss Journal of Economics and Statistics*, Vol. 116/III, pp. 235-260.
- Kimura, F. (2006), "International production and distribution networks in East Asia: Eighteen facts, mechanics, and policy implications", *Asian Economic Policy Review*, Vol. 1/2, pp. 326-344.
- Kohler, W. (2004), "Aspects of International Fragmentation", *Review of International Economics*, Vol. 12/5, pp. 793-816.
- Koopman, R. and S.-J. Wei (2012), "The Value-added Structure of Gross Exports: Measuring Revealed Comparative Advantage by Domestic Content in Exports", working paper.
- Koopman, R. Z. Wang and S.-J. Wei (2011), "Give credit to where credit is due: tracing value added in global production chains", *NBER Working Paper 16426*, September 2010, revised September 2011.
- Krishna, P. and D. Mitra (2008), "Reciprocated unilateralism in trade reforms with majority voting", *Journal of Development Economics*, Vol. 85/1, pp. 81-93.
- Krishna, P. (2013), "Preferential Trade Agreements and the World Trade System: A Multilateralist View", in Feenstra and Taylor (eds.), *Globalization in an Age of Crisis: Multilateral Economic Cooperation in the Twenty-First Century*, University of Chicago Press. Also *NBER Working Paper 17840*.

- Krugman, P. (1991), "Increasing Returns and Economic Geography", *Journal of Political Economy*, Vol. 99, pp. 483-99.
- Krugman, P. and A.J. Venables (1995), "Globalization and the Inequality of Nations", *The Quarterly Journal of Economics*, Vol. 110/4, pp. 857-80.
- Krugman, P. (1991), "The Move Toward Free Trade Zones", in "Policy Implications of Trade and Currency Zones, A Symposium Sponsored by The Federal Reserve Bank of Kansas City", Jackson Hole, Wyoming, 22-24 August.
- Kurz, H. and N. Salvadori (2006), "Input–Output Analysis from a Wider Perspective: a Comparison of the Early Works of Leontief and Sraffa", *Economic Systems Research*, Vol. 18/4, pp. 373-3906.
- Lamy, P. (2010), "Globalization of the industrial production chains and measuring international trade in value added", Speech to French Senate, 15 October 2010.
- Lanz, R., S. Miroudot and H.K. Nordås (2012), "Does fragmentation of production imply fragmentation of jobs?", GTAP Conference paper, https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=3916
- López González, J. (2012), "Vertical Specialisation and New Regionalism", PhD thesis, University of Sussex, April 2012.
- Ludema, R., A. Mayda and P. Mishra (2010), *Protection for Free?: The Political Economy of US Tariff Suspensions*, Centre for Economic Policy Research, Washington DC.
- Lupo Pasini, F. (2011), "The International Regulatory Regime on Capital Flows", *ADB Working Paper*, No. 338, Asian Development Bank, 30 December 2011.
- Mattoo, A. and C. Fink (2004), "Regional Agreements and Trade in Services: Policy Issues", *Journal of Economic Integration*, Center for Economic Integration, Vol. 19, pp. 742-779.
- Mayer, W. (1985), "The political economy of tariff agreements", *Schriften des Veriens fur Sozialpolitik*, Vol. 148, pp. 428-437.
- Miroudot, S., J. Sauvage and M. Sudreau (2010), "Multilateralising Regionalism: How Preferential Are Services Commitments in Regional Trade Agreements?", *OECD Trade Policy Papers*, No. 106.
- Moser, P. (1990), *The political economy of the GATT*, Verlag Ruegger, St Gallen.
- Neumayer, E. and L. Spess (2005), "Do Bilateral Investment Treaties Increase Foreign Direct Investment to Developing Countries?", *World Development*, Vol. 33/10, pp. 1567-1585.
- Niepmann, F. and G. Felbermayr (2010), "Globalisation and the Spatial Concentration of Production", *The World Economy*, Vol. 33/5, pp. 680-709.
- O'Rourke, K.H. & J.G. Williamson (2002), "When did globalisation begin?," *European Review of Economic History*, Vol. 6/01, pp. 23-50.
- Orefice, G. and N. Rocha (2013), "Deep integration and production networks: an empirical analysis", forthcoming, *The World Economy*.
- Page, S. and P. Kleen (2005), "Special and Differential Treatment of Developing Countries in the World Trade Organization", *Global Development Studies*, No. 2, Swedish Ministry for Foreign Affairs, Sweden.
- Patterson, G. (1966), *Discrimination in international trade: The policy issues 1945-1965*, Princeton University Press.
- Porter, M.E. (1985), *Competitive Advantage*, Free Press, New York.
- Pritchett, L. (1997), "Divergence, Big Time", *The Journal of Economic Perspectives*, Vol. 11/3, pp. 3-17.
- Puga, D. and A.J. Venables (1996), "The Spread of Industry: Spatial Agglomeration in Economic Development", *Journal of the Japanese and International Economies*, Vol. 10/4, pp. 440-464.
- Roessler, F. (1978), "The Rationale for Reciprocity in Trade Negotiations under Floating Currencies", *Kyklos*, Vol. 31/2, pp. 258-274.

- Sanyal, K., and R. Jones (1982), "The theory of trade in middle products", *American Economic Review*, Vol. 72/1, pp. 16-31.
- Simchi-Levi, D. (2010), "Impact of Crude Oil Volatility on Network Design", online Powerpoint presentation, www.scdigest.com/assets/rep/Impact_Oil_Prices_on_Supply_Chain_Network_Design.pdf.
- Simchi-Levi, D. et al. (2011), "Made in America: Rethinking the Future of US Manufacturing", Accenture, www.accenture.com/SiteCollectionDocuments/PDF/Accenture-Made-in-America.pdf.
- Spitz, A. (2004), "Are Skill Requirements in the Workplace Rising? Stylized Facts and Evidence on Skill-Biased Technological Change", *Discussion Paper No. 04-33*, Centre for European Economic Research, Mannheim.
- Swenson, D.L. (2005), "Overseas assembly and country sourcing choices", *Journal of International Economics*, Vol. 66/1, pp. 107-130.
- Timmer, M. (ed.) (2012), "The World Input-Output Database (WIOD): Contents, Sources and Methods", *WIOD Working Paper 10*, <http://www.wiod.org/publications/papers/wiod10.pdf>
- Venables, A. J. (1999), "Fragmentation and multinational production", *European Economic Review*, Vol. 43, pp. 935-945.
- Vézina, P.-L. (2010), "Race-To-The-Bottom Tariff Cutting", *IHEID Working Papers*, No. 12-2010, Economics Section, Graduate Institute of International Studies, revised July 2010.
- Woodland, A. (1977), "Joint outputs, intermediate inputs and international trade theory", *International Economic Review*, Vol. 18/3, pp. 517-533.
- WTO (2011), "The WTO and reciprocal preferential trading agreements", *World Trade Report 2011*, Geneva.
- Yeats, A.J. (1998), "Just How Big is Global Production Sharing?", *Policy Research Working Papers*, No. 1871, The World Bank, Washington DC.
- Yi, K.M. (2003), "Can vertical specialization explain the growth of world trade?", *Journal of Political Economy*, Vol. 111/1, pp. 52-102.