Centre for Economic Policy Research
POLICY INSIGHT No.56
MAY 2011

21st Century Regionalism: Filling the gap between 21st century trade and 20th century trade rules

Richard Baldwin

Graduate Institute, Geneva and CEPR

1. Introduction

The last time multilateral trade rules were updated, Bill Clinton was in his first term of office, data was shared by airmailing 1.4 megabyte HD floppy disks (few people had email), cellphones looked like bricks and calling costs were measured in dollars per minute. Trade mostly meant selling goods made in a factory in one nation to a customer in another. Simple trade needed simple rules – a fact reflected in both multilateral and regional trade agreements.

Today's trade is radically more complex. The ICT revolution fostered an internationalisation of supply chains, and this in turned created the "trade-investment-service nexus" at the heart of so much of today's international commerce. Complex trade needs complex rules. As the WTO was otherwise occupied, the incipient governance gap was filled by uncoordinated developments elsewhere – primarily in deep regional trade agreements, bilateral investment treaties, and autonomous reforms in emerging economies. The resulting package of deeper disciplines – what could be called "21st century regionalism" – requires new thinking.

This paper marshals several sets of facts into an argument that:

- Today regionalism is qualitatively different to that of the 1990s;
- The traditional building-stumbling-block approach and Vinerian economics on which it is premised are not up to the job of analysing this new regionalism; and
- 21st century regionalism has quite different ramifications for the world trading system than 20th century regionalism did.

The facts are not new, having been documented by many observers. The aim of the paper is to weave them together into an argument that points to a new way of thinking about regionalism – specifically, its economic implications, its political economy determinants, and its impact on the world trade system and the WTO.

In a nutshell, 21st century regionalism is not primarily about preferential market access as was the case for 20th century regionalism; it is about disciplines that underpin the trade-investment-service nexus. This means that 21st century regionalism is driven by a different set of political economy forces; the basic bargain is "foreign factories for domestic reforms" – not "exchange of market access". As 21st century regionalism is largely about regulation rather than tariffs, regulatory economics is needed rather than Vinerian tax economics. Finally, 21st century regionalism is a serious threat to the WTO's centrality in global trade governance, but not for the reason suggested by the old building-stumbling-block thinking. 21st century regionalism is a threat to the WTO's role as a rule writer, not as a tariff cutter.

Editors note: The paper was first presented at the WTO's "Workshop on PTAs and the WTO: A new era" held at the WTO 4 November 2010; http://www.wto.org/english/res_e/reser_e/wkshop_nov10_e/wkshop_nov10_e.htm. My thanks to comments from the participants as well as seminar participants at Keio University and RIETI in Tokyo, Georgetown University, the Indian Institute for Foreign Trade in Delhi, and the Taiwan WTO Center (Chung-Hua Institution for Economic Research), and comments from Theresa Carpenter, Simon Evenett, Patrick Low, Ted Moran, and Rod Ludema. Special thanks to Alan Maldic, Yose Damuri, Andy Lendle for help with data analysis.

Plan of the paper

The next section discusses the complexity of 21st century commerce, how it arose and why. The subsequent section, Section 3, discusses how this more complex trade created a demand for more complex disciplines and how 21st century regionalism met these demands. Having laid out the problem, Section 4 argues that a new analytic framework is needed to think about 21st century regionalism. Section 5 discusses features the new framework should display. Section 6 considers the implication for the world trading system and presents some concluding remarks.

2. Emergence of 21st century trade

Today's international commerce comprises complex, two-way flows of goods, services, people, ideas, and investments in physical, human and knowledge capital – in addition to trade in raw materials and final goods. There is nothing new about this from a qualitative sense. It has been going on for decades. The 1957 Treaty of Rome and the 1965 US-Canada Auto Pact, for example, were designed to encourage just such exchanges (Wonnacott, 1987).

The novelty of 21st trade lies in the quantitative dimension. To illustrate what is really new, it is useful to put 21st century trade into the broad context of globalisation.

2.1 Globalisation as two unbundlings

When sailing ships and horse carts were state-of-the-art transportation, only items with very high value-to-weight ratios could be profitability shipped over anything but the shortest distances. As a result, each village made most of what it consumed; production and consumption were geographically bundled. Steam power changed this (O'Rourke and Williamson 1999, Chapter 3).

Railroads and steamships radically lowered transport costs, so production and consumption could be unbundled geographically. Once unbundling was feasible, scale economies and comparative advantage made it inevitable. This was globalisation's "first unbundling" although it occurred in two waves punctuated by world wars and the Great Depression.1

Global unbundling with local clustering

Globalisation's first unbundling generated a paradox. As production dispersed internationally, it clustered locally (into factories and industrial districts). Better transportation favoured scale economies that typically involved complex manufacturing processes. This new complexity fostered local clustering. For example, consider a stylised production process with several production bays. Coordination involves a continuous, two-way flow among the bays of good, people, ideas, and investment in machines, training and technology. The two-way flows never cease as continuous efforts to heighten productivity keep the process in flux. Production clustered locally because proximity lowered the cost of the two-way flows. A new distance-linked cost became important – what might be called "coordination glue".

ICT revolution: 2nd unbundling's equivalent of the steam revolution

Some of this coordination glue is related to communication. As telecommunications became cheaper and surer from the mid-1980s, the coordination glue began to weaken. The price of telephone calls plummeted, faxes became standard, cellular phone usage exploded, and the telecommunication network became denser, more reliable and cheaper. Two other trends interacted with cheaper communication costs – the spectacular fall in the price of computing power (Moore's Law) and the equally spectacular rise in fibre optic transmission rates (Gilder's Law). Long-distance information sharing was revolutionised as these developments in telecoms were complemented by the rise of the internet – first email and then web-based platforms.

The telecom and internet revolutions triggered a suite of information-management innovations that made it easier, cheaper, faster, and safer to coordinate complex activities at distance. Email, editable files (*.xls, *.doc, etc), and more specialised web-based coordination software packages revolutionised peoples' ability to manage multifaceted procedures across great distances. Working methods and product designs also shifted to make production more modular and thus easier to coordinate at distance. Stages of production that previously had to be performed in close proximity – within walking distance to facilitate face-to-face coordinate of innumerable small glitches – could now be dispersed without an enormous drop in efficiency or timeliness. Collectively, this is known as the information and communication technology (ICT) revolution.

Figure 1, which displays several ICT, indicators, shows that there was an inflection point in the growth of internet hosts in 1985 and in telephone subscribers in 1995. This suggests that the coordination glue began to weaken sometime between 1985 and 1995.

1 For estimates of trade flows back to 1870, see David, Meissner, and Novy (2011). For a detailed account of the two waves of the first unbundling, see Baldwin and Martin (1999).

Figure 1. Growth of global internet hosts and phone lines, 1975 – 2011.

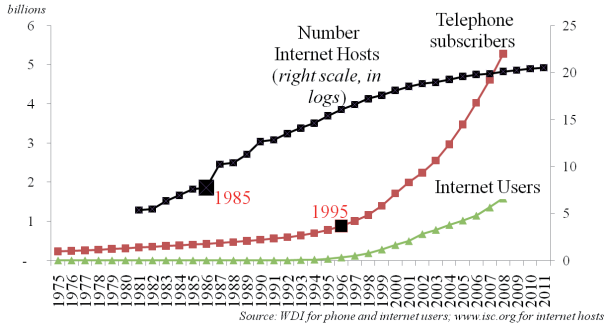
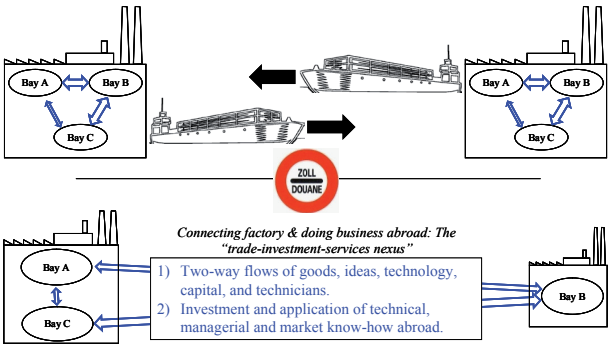


Figure 2. Schematic illustration of 20th and 21st century trade



As far as trade is concerned, the partial melting of the coordination glue meant that some production stages that previously had to be within walking distance could now be dispersed internationally. Once ICT made dispersion of production stages feasible, scale economies and comparative advantage made it inevitable. This is globalisation's "second unbundling" – the spatial unbundling of productions stages previously clustered in factories and offices.[2] This radically changed the nature of international commerce giving rise to what might be called the trade-investment-service nexus.

2.2 The trade-investment-services nexus: 21th century trade

The 2nd unbundling did not end the need to coordinate production stages – it internationalised it. International commerce became more complex. The result might be called 21st century trade. The heart of 21st century trade is an intertwining of: 1) trade in goods, 2) international investment in production facilities, training, technology and long-term business relationships, and 3) the use of infrastructure services to coordinate the dispersed production, especially services such as telecoms, internet, express parcel delivery, air cargo, trade-related finance, customs clearance services, etc. This could be called the trade-investment-services nexus.

The differences between 20th and 21st century trade are illustrated schematically in Figure 2. The top panel illustrates 20th century trade; trade is dominated by goods made in factories in one nation and sold to customers in another. There are complex two-way flows of goods, people, and ideas (the double-headed arrows) but primarily within factories. The lower panel illustrates 21st century trade. Here factories and offices have been unbundled internationally thus creating the trade-investment-service nexus where some of the complex two-way flows that used to take place within factories and offices now take place across international borders.

It is useful to think of the trade-investment-services nexus as being created by two distinct sets necessities:

- Connecting factories, and
- Doing business abroad.

Trade in parts and components, trade in infrastructure services, and foreign direct investment are the most easily measured aspect of this multifaceted, multi-directional commerce, but they are only the tip of the proverbial iceberg.

2.1 Indicators of the second unbundling

The "second unbundling" happened first among developed nations. Examples include US-Canada or French-German trade in autos and auto parts in the 1970s (Hummels, Rapoport, and Yi, 1998). The big change, however, came when the second unbundling accelerated between developed and developing nations (Hanson and Feenstra 1997, and Ando and Kimura 2005). The dominant factor here was the juxtaposition of the ICT revolution and colossal wage discrepancies.[3]

Many scholars have documented the nature of this 'new' trade. For example, one very obvious form of production unbundling is known as outward processing trade, or vertical specialisation trade as documented by Ishii and Yi (1997). Here intermediate inputs are imported and used in goods that are subsequently exported. Amador and Cabral (2008) show that this trade was more important in among European and North American nations up until the early to mid 1980s. After that, it boomed North-South, but especially in Asia (Figure 3).

Internationalisation of the supply chain however is a much broader phenomenon than outward

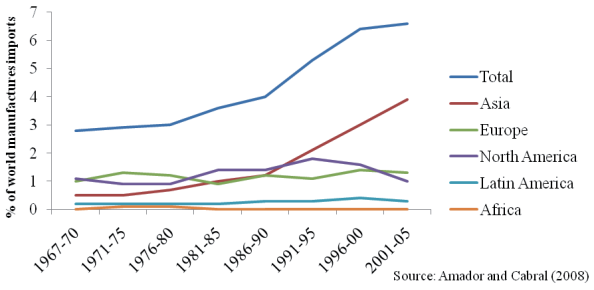
2 See Baldwin (2006a) for the original presentation of globalisation as two unbundlings and policy implications for European social welfare states.

3 For firm-level evidence, see Ariu and Mion (2010).

processing trade.[4] In the case of East Asia, production unbundling has progressed to a state where the term "Factory Asia" is widely used.

Development of Factory Asia can be tracked with the Asian Input-Output table maintained by Japan's JETRO.[5] In 1985, Factory Asia was simple. The manufacturing sectors of developing East Asia – which meant all nations other than Japan – supplied most of their own intermediates. And what they did not supply themselves, they imported from technologically advanced nations, mostly Japan, the US and the EU. There was very little trade in intermediates among them. By the 1990s, local sourcing of parts and components fell, matched by an increase in imports from Japan, the US, Korea, Taipei China, Singapore, and Hong Kong.

Figure 3 Outward processing trade, 1967 – 2005..

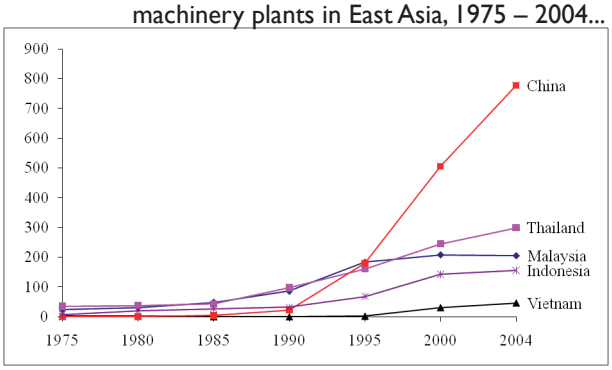


Moving ahead to 1995 and 2000, the trends generally continue, but by 2000, intermediate trade among developing East Asian nations had become a significant phenomenon. Countries like Thailand and China were no longer simply engaged in outward processing trade. They were supplying considerable shares of intermediate inputs into the manufacturing sectors of their fellow developing nations. Japan's and China's share of own-intermediates, however, hardly budged during this time.

Evidence for these trends on a global scale can be found in a recent important paper, Johnson and Noguera (2010). They document just how far supply chains have been internationalised.

Another important – and easily measured – facet of supply-chain internationalisation is foreign direct investment. This also flourished at approximately the same time, namely the mid-1980s and early 1990s. Figure 4 illustrates the case of Japanese auto and electrical machinery plants placed in East Asian nations. The evolution shows a clear acceleration from 1985 with another inflection point in the mid-1990s – mostly due to plants placed in China.

Figure 4 Number of Japanese auto and electrical machinery plants in East Asia, 1975 – 2004...



Source: Fujita and Hamaguchi (2006).

3. Filling the governance gap: 21st century regionalism

Simple commerce needs simple rules; complex commerce needs complex rules. When trade meant factories in one nation selling goods to customers in another, international rules could be simple – dealing primarily with border measures and a few 'behind-the-border barriers' (BBBs) such as blatantly discriminatory national taxes and regulations. Multilaterally, the 1947 GATT embodied this simple set of disciplines. 20th century regional trade agreements (RTAs) were even 'shallower' – frequently addressing little more than tariffs and rules of origin.

As 21st century trade became more complex, demands arose for more complex international trade rules. This section argues that this demand was met by collection of uncoordinated developments outside the WTO – a collection that might be called 21st century regionalism. The nature of these new demands is discussed before considering how 21st century regionalism filled the governance gap.

3.1 Demands for deeper economic integration rules

When it comes to governance, the critical difference between 20th and 21st century trade is the trade-investment-services nexus. As discussed above, the nexus entails two elements, each of which generated new demands for more complex international disciplines:

- Doing business abroad.

When firms set up production facilities abroad – or form long-term ties with foreign suppliers – they typically expose their capital as well as their technical, managerial and marketing know-how

4 See, for example, Ando and Kimura (2005), Kimura, Takahashi, and Hayakawa (2007), Gaulier, Lemoine and Unal-Kesenci (2007), and Athukorala (2005) in the East Asian case, and Dallas Fed (2002) or Feenstra and Hanson (1996) on the North American case.

5 See Baldwin (2006b) for a more detailed analysis of Factory Asia using the Asian IO matrices.

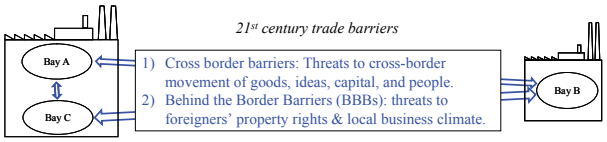
to new international risks.[6] Threats to these tangible and intangible property rights became 21st century trade barriers.

- Connecting international production facilities.

Bringing high-quality, competitively-priced goods to customers in a timely manner requires international coordination of production facilities via the continuous two-way flow of goods, people, ideas and investments.[7] Threats to these flows became 21st century trade barriers.

Figure 5 illustrates 21st century trade barriers schematically. Here are some examples.

Figure 5 The disciplines to underpin the "trade-investment-services nexus".



- The sharing of tacit and explicit technology and intellectual property is facilitated by assurances that foreign knowledge-capital owners will be treated fairly and their property rights will be respected.

- Foreign investments in the training of workers and managers, physical plant, and the development of long-term business relationships are facilitated by assurances on property rights, rights of establishment, and anticompetitive practices.

- Assurances on business related capital flows – ranging from new FDI to profit repatriation – also helped foster the investment part of the trade-investment-services nexus.

- Connecting factories often involves time-sensitive shipping, world class telecoms and short-term movement of managers and technicians, so assurances on infrastructure services are also important.

- Tariffs and other border measures also matter – just as they mattered in the 20th century but more so since the ratio of value added to value on individual shipment falls as the production chain fragments, even though tariffs are applied to the value of the goods as they cross borders.

This list suggests four types of 21st century trade barriers that were not barriers to 20th century trade: competition policy (known as anti-trust in the US), movement of capital, intellectual property rights beyond the TRIPs Agreement, and investment assurances.

Supply meets demand

21st century trade was welcomed by many nations. To encourage it, they wanted to remove 21st century trade barriers. This is the ultimate source of demand for more complex international disciplines. Multinationals from advanced-technology nations were eager to lower production costs by dispersing production and technology to the most cost-effective locations. Developing-nation governments embraced 21st century trade as a fast lane to industrialisation and growth. In short, the mutual interest in 21st century trade meant that the rising demands for new discipline were met by willing suppliers – namely governments.

This complex trade first arose among rich nations (Figure 3). These nations' high levels of property rights, legal transparency and availability of infrastructure services helped underpin 21st century trade even without WTO or RTA disciplines. The lack of deeper disciplines became a first-order problem when 21st century trade started to involve emerging economies with weaker domestic governance.

This is how 21st century regionalism first emerged. Emerging market economies– eager for advanced technology factories – were willing to embrace disciplines on things that were not traditionally considered to be trade barriers. As the WTO was otherwise occupied, the demand for deeper discipline was filled by:

- Deep RTAs,

- Bilateral Investment Treaties (BITs), and

- Unilateral reforms.

Consider these in turn.

6 As World Bank (2011) notes doing business abroad implicates "the laws, regulations and institutional arrangements that shape daily economic activity." The 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.

7 Tariffs on imported intermediates are one part of this. Coordinating international production also requires assurances of world-class telecommunications, world-class goods transportation (especially express parcel services and air cargo) and customs clearance, transportation and assured access for short-term visits by key personal (managers and technicians), and capital and financial market openness to inward and outward investment flows and profit repatriation.

3.2 Deep RTAs

The distinction between deep and shallow RTAs was highlighted explicitly by Lawrence (1996). He also noted its association with more complex trade and pointed out that it first developed among developed nations in Europe and North America. From the mid-1990s, deep RTAs spread rapidly to cover North-South trade. The US-Mexico component of NAFTA and Europe's Euro-Med Association Agreements led the way. More recently, Japan has joined the movement by signing deep Economic Partnership Agreements (EPAs) with the large ASEAN economies. Note, however, that even South-South agreements have deepened in the 21st century.

Table 1 WTOx and WTO+ provisions in regional trade agreements.

	Pre-WTO	1995-2000	DDA era, post 2001
		WTO+ Issues	
Customs	13	11	56
AD	12	8	53
CVM	4	5	52
Export Taxes	8	8	41
State Aid	10	9	34
TRIPs	6	4	41
GATS	7	2	39
STE	5	3	35
TBT	2	2	36
SPS	2	1	35
Public Procurement	5	0	32
TRIMs	6	2	31
		WTOx Issues	
Competition Policy	11	9	39
Movement of Capital	6	5	38
IPR	5	2	39
Investment	4	1	35

Source: World Trade Report, WTO (2011).The WTO+ and WTOx classification is from Horn, Mavroidis and Sapir (2010).

Recent research has started quantifying the depth of various RTAs based on the methodology of Horn, Mavorodis and Sapir (2010). These authors read through all the US and EU agreements and noted whether they contained WTO+ and WTOx provisions, and whether these provisions were legally enforceable. 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. In all Horn and co-authors identify 14 WTO+ and 38 WTO-X policy areas. Yap, Medalla and Aldaba (2006) and Balboa (2008) have done a similar exercise on Japanese EPAs.

The 2011 World Trade Report by the WTO Secretariat extends the Horn-Mavroidis-Sapir method to 97 more RTAs. Of this, 33 involved the EU and 11 involved the US. The 43 other RTAs were concluded by regional trading blocs and major trading powers such as ASEAN, China and MERCOSUR. The sample of RTAs was chosen based primarily on their volume of intra-RTA trade. The agreements investigated came into force during the 1958 to 2010 period.

Using their new dataset, WTO (2011) concludes that RTAs have been getting increasingly 'deep' with most of the provision being legally enforceable. As they note: "The pattern observed suggests that deepening commitments in these areas, i.e., going beyond commitments in the WTO, continue to be a major driving force for recent RTAs." The report also points out that RTAs in the 21st century cover more WTO-X areas than earlier RTAs. The main policy areas covered are competition policy, intellectual property rights, investment and movement of capital. As these WTO-X provisions are largely regulatory in nature, their growth is "testimony to the growing importance of behind-the-border measures in RTAs."

One of the most striking results from the WTO's new work is the identification of a core of four deeper disciplines that appear in over a third of RTAs, but which are not part of the WTO's rulebook. These are competition policy (47% of all agreements), movement of capital (39%), intellectual property rights not in the TRIPs Agreement (37%), and investment (31%).

Deep RTAs are also defined by what they are not. As it turns out, tariff preferences are no longer very important at a global scale, although they are still important in certain sectors and certain RTAs.

3.2.1 RTAs are no longer very preferential

While bilateral trade flows covered by RTAs account for about half of world imports, only 16.7% of world trade is eligible for preferences (Carpenter and Lendle 2010). The remaining trade flows either have zero MFN tariffs (about 25% of world trade) so there can be no preference, or they are excluded from preferential treatment by the terms of the RTA (about 9% of world trade). In fact, the largest traders have lowered their applied MFN tariffs to zero on a wide swath of imports. The figure for the EU's external trade is 56%, while it is 43% for the US, 48% for China, and 80% for Japan. The products where the large importers maintain high tariff are routinely excluded from their RTAs, so again no preference arises. A second important fact is that the margins of preferences are quite low on the 16.7% of world trade that is preferential. In fact, less than 2% of world imports enjoy preferences over 10 percentage points.[8]

8 These numbers are calculated without considering intra-EU trade. Taking world totals (including intra-EU flows) Carpenter and Lendle (2010) calculate that 64% of world trade is covered by an RTA, 29.8% of world trade is subject to preference margins, but only 3.9% of this enjoys margins over 10 percentage points.

Table 2. Margins of preference in 2008.

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

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

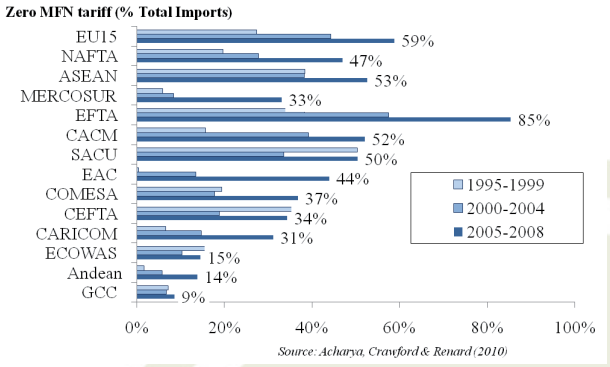
As Table 2 shows, imports of the largest importers are not subject to huge preference margins. Intra-EU trade is by far the most preferential with 9% carrying preference margins over 10 percentage points. Preferences in this range granted by the US amount to 2% of its imports; China and Japan grant such preference margins to essentially none of their imports.

The nations with the highest share of imports that are covered by large preferences (i.e. over 10% margins) are small nations that are heavily dependent on large neighbours. For instance, 16% of Mexican and 2% of Canadian imports receive over 10% margins, while the corresponding number for Turkey is 2%. For these nations, and Chile, an important share of imports received margins of between 5 and 10 percentage points.

The massive data demand for the calculations in Table 2 prevent Carpenter and Lendle (2010) from covering all plurilateral RTAs beyond the EU and NAFTA. A rougher set of facts, however, is available for a broader range of RTAs. Archarya, Crawford and Renard (2010) calculated the share of various RTAs imports that are subject to zero MFN applied rates and thus not eligible, de facto, for preferences.

The results, shown in Figure 6, confirm the basic fact that many so-called preferential trade agreements are not in fact very preferential. The major North-North RTAs all have half or more of their imports duty-free on an MNF basis. The developing nation RTAs, especially in African and Latin America have much lower shares of MFN duty-free imports (the South African Customs Union, SACU, and the Central American Common Market are exceptions), but in most cases, the share of MFN duty-free imports has increased significantly since 1995 – with the increase for Mercosur being particularly noticeable.

Figure 6 Share of imports with MFN zero tariffs, various RTAs, 1995 to 2008..



Another excellent study illustrating the same basic facts is Fugazza and Nicita (2010). They go one step further in considering interactions between preference margins and import elasticities. The idea is to check whether the large preferences fall on goods where quantity responds abundantly to small price differences. Despite this refinement, they come to the same conclusion that tariff preferences are now rather small from a global perspective.

Deep RTAs are not the only non-WTO route to establishing the disciplines required for 21st century trade. Equally important has been the explosion of Bilateral Investment Treaties.

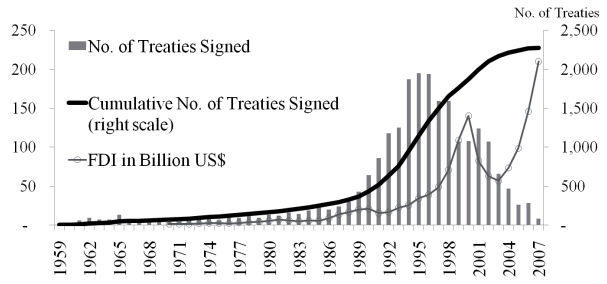
3.3 Bilateral Investment Treaties

BITs have long attracted the attention of trade scholars, although much more from trade lawyers than from trade economists (Sauvant and Sachs 2009).[9] Bilateral Investment Treaties (BITs) establish disciplines that govern interactions between private foreign investors and host governments. As such, they are central to the investment part of the trade-investment-services nexus which forms the core of 21st century trade.

BITs are extremely common with about 2,500 in existence – almost five times the number of RTAs. This number, however, overstates the coverage. Investment policy – unlike trade policy – is not in the hands of EU but rather individual EU members, so, for example, 23 of Turkey's 82 BITs are with EU members. Overall, more than 1,200 BITs have been negotiated by individual EU members (TNI 2011).

As Figure 7 shows, the explosion of BITs coincided with the second unbundling (which took off between 1985 and 1995). As international commerce became more complex – in particular as the investment part of the trade-investment-services nexus developed, FDI emitters signed BITs with FDI seekers. There may well have been a sort of domino or 'race to the bottom' element in the BIT proliferation, as FDI seekers felt they had to sign BITs to guard their locational competitiveness against rivals who had already signed BITs with the major source-nations of FDI.

Figure 7 Explosion of Bilateral Investment Treaties...



Source: UNCTAD's World Investment Report, various issues

The basic aim of BITs is to encourage FDI, often FDI related to the trade-investment-services nexus. All the major FDI emitters – the Europeans, the US, Japan, etc. – have their own model agreements. The US model is explicit and quite comprehensive, so it serves well as an illustration of the basic features of a BIT.[10]

The six basic goals of a typical US BIT are:

To assure non-discrimination. Investments abroad are to be treated as favourably as the host party treats its own investors (national treatment) and third country investors (most favoured nation).

To discipline expropriation. Expropriations are to be limited and when they occur, they must be compensated by prompt and adequate payment.

To assure transferability of investment-related funds. Capital and financial markets are to be open so as to permit investment-related capital inflows and outflows without delay and at market rate of exchanges.

To discipline the imposition of performance requirements. Host government policies that tie the hands of investors – such as local content targets or export quotas, as a condition for the establishment, acquisition, expansion, management, conduct, or operation of an investment – are to be limited.

To assure investors ability to choose top managers. Host-government policies relating to the choice or nationality of top management is to be limited, or prohibited.

To grant private investors the right to submit disputes with the host government to international arbitration as opposed to local courts. This feature is one of the most distinctive and it has lead to construction of a new jurisprudence completely independent of the WTO's dispute settlement procedure. The main "court" is the International Centre for Settlement of Investment Disputes (ICSID) based in Washington.

Not just North-South

Up to 1980, almost all BITs involved at least one old OECD nation, but BITs among OECD nations were rare. For the most part, BITs were North-South agreements. Starting in the 1980s, with the trend accelerating in the 1990s, new OECD members (see the table note for a list) and non-OECD nations began signing them. BITs are no longer only North-South agreements. The second unbundling has seen a proliferation of South-South FDI which also benefitted from BIT-disciplines and assurances.

9 BITs impact on FDI flows has been recently estimated by, for example, Egger and Merlo (2007).

10 See, for example, http://www.state.gov/documents/organization/117601.pdf.

This development is very much in line with the data from Japan's Asian Input-Output data discussed above. In the 1980s, the supply chain was fairly simple – what was known as "triangle trade" in Asia (METI 2005 Section3) – where Japan and the NIEs produce intermediate goods, China and ASEAN import intermediate goods, assemble them into final goods, and then export to the US and EU. As the second unbundling proceeded, the linkages in Factory Asia became denser. The trade-investment-service nexus developed among developing nations and thus arose a demand for the property assurances in BITs.

Table 3 shows the facts. The five panels in the table each show a decade starting in 1959. The figures reflect that number of BITs signed in the relevant decade between partners in the three groups – old OECD, new OECD and non-OECD. In the 1990s, almost 500 BITs were signed among non-OECD nations. The BIT network is especially dense among the leading East Asian emerging economies: China, Thailand, Indonesia, Malaysia and the Philippines. There is another BIT network among the Andean nations.

As far as the old OECD nations are concerned, the flow of new BITs has tapered off significantly, but South-South BITs continue to flourish. The world leader in BITs is Germany with 147; the US is in 33rd place with "only" 48.

Table 3 Number of BITs by decade and by partner pairings.

		Nation 2		
	Nation 1	Old OECD Members	New OECD Members	Non-OECD
1959-69	Old OECD Members	1		
	New OECD Members	2		
	Non-OECD	62		
1970-79	Old OECD Members			
	New OECD Members	11		
	Non-OECD	67	1	9
1980-89	Old OECD Members	1		
	New OECD Members	30	2	
	Non-OECD	141	13	28
1990-99	Old OECD Members			
	New OECD Members	86	15	
	Non-OECD	606	202	485
2000-07	Old OECD Members			
	New OECD Members	12	3	
	Non-OECD	234	41	224

Source: Database of BITs compiled by ICSID http://icsid.worldbank.org/ICSID/FrontServlet.

Notes: Old OECD: Australia, Austria, Belgium-Luxembourg, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK, US. New OECD: Czech Republic, Hong Kong, Hungary, Korea, Mexico, Poland, Singapore, Slovak Republic, Turkey.

The 1990s explosion of BITs had two margins of expansion. The number of BITs signed per nation increased, and the number of nations signing BITs increased. Before the second unbundling started in the late 1980s, BITs were initially mainly between European FDI emitters and developing nation FDI seekers. The UK and Japan signed their first treaties in the 1970s and the US only in the 1980s. The list of developing country BIT signers expanded rapidly in the 1990s and since many of them signed BITs with the major FDI emitters (the big EU nations, the US and Japan) as well as with other developing nations in their region, the number of new BITs in the 1990s top a thousand. By the end of the 1990s, almost all WTO members were playing the BITs game.

3.4 Unilateralism

Many developing nations actively seek the participation of foreign companies in their economies and indeed set a central role for such participation in their industrialisation strategies. As part of this, many have signed the deep RTAs and BITs as discussed above. Even more, however, have embarked on unilateral policy reforms that improve their business climates, especially for foreign firms in manufacturing or trade-oriented services. It also improves the competitiveness of local manufacturing by lowering the cost of imported parts and components.

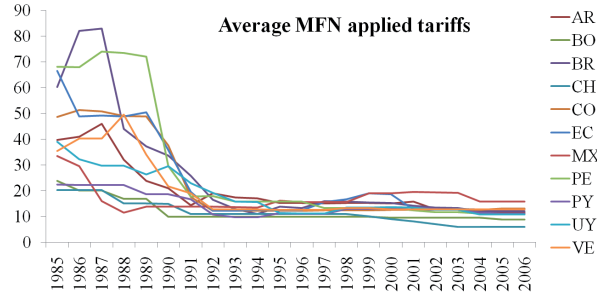
Evidence for the broad reforms is difficult to marshal for many nations going back to the early days of the second unbundling (1985-1995). One aspect, however, is easily measured – tariffs. The spectacular tariff reductions in Latin America are shown in Figure 8. Clearly there is substantial variation among nations, but a clear general towards unilateral tariff reduction from the late 1980s. Note that now most Latin American tariff averages are around 10%. It should be noted that Latin Americans where cutting tariffs in RTAs at approximately the same rate as they were lowering their MFN applied rates.

Unilateralism has also proceeded in other developing nations. Some of this, especially in Africa, was driven by IMF conditionality, but even nations not under such external pressure lowered rates, as Figure 9 shows.

According to new evidence presented in WTO (2011), the global reduction in tariffs on parts and components has exceeded the overall average. This is rough evidence of an association between the 2nd unbundling and autonomous tariff liberalisation.[11]

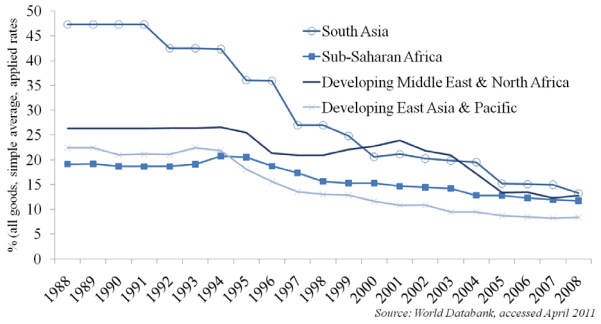
11 On the political economy of unilateralism, see Garnaut 1991, Young 1996, Edwards and Lederman 1998, Richardson 2001, and Sally 2008, Coates and Ludema (2001), Krishna and Mitra (2008), and very recently Ludema, Mayda and Mishra (2010), Conconi and Perroni (2010), and Baldwin (2010).

Figure 8 Applied MFN tariff liberalisation in Latin America.



Source: IADB and ITC database.

Figure 9 Average tariffs, South Asia, East Asia, Middle East & North Africa and Sub-Saharan Africa



4. New analytic framework needed

Plainly regionalism has changed dramatically. How should we think about it? Can one continue to use the traditional approach to regionalism that was established two decades ago to understand shallow RTAs and 20th century trade?

In the early 1990s, regionalism metastasised just as multilateral GATT talks slipped into a four-year coma. These facts mandated an analytic approach that focused on regionalism versus multilateralism. Krugman (1991a,b) Bhagwati (1991, 1993), and Summers (1991) were the architects of the new approach, but it has come to be known by Jagdish Bhagwati's bon mot "building blocks versus stumbling blocks". Two assumptions form the pillars of the building-stumbling-block logic:

- It is useful to think of RTAs as being exclusively about preferential tariffs (non-tariff measures are unimportant, or can be modelled as tariffs).[12]

The identification of RTAs with tariff preferences (sometimes called the PTA-view) meant that traditional thinking was based on Vinerian economics.[13]

- It is useful to frame the trade-off as regionalism-versus-multilateralism (unilateralism can be ignored or folded into multilateralism).[14] In evaluating the tradeoffs, it is useful to take regionalism as exogenous but the impact on multilateralism as endogenous.[15]

Consider the evidence on whether these two pillars are valid for 21st century regionalism.

4.1 Tariff preferences and Vinerian economics

Section 3.2.1 showed that it is factually incorrect to view RTAs as being mainly about preferential tariffs. This directly suggests that the PTA-view of RTAs falls down when it comes to 21st century regionalism. The reliance on Vinerian economics, however, could be defended if the deep provisions in RTAs mimicked preferential tariffs. As we shall see, some non-tariff measures in RTAs do act like preferential tariffs, but many do not. This assertion can be demonstrated with empirical evidence and economic logic. To interpret the evidence correctly, however, it is useful to review Vinerian logic.

Vinerian economics is nothing more than tax analysis when different rates are applied to

12 All the pioneering works of Bhagwati, Krugman and Summers focused on tariffs, and virtually all the theory articles stimulated by this early work assume they are only about preferential tariffs. Among the best-known are Aghion, Antràs, Helpman (2007), Bond and Syropolous (1996), Freund (2000a,b), Grossman and Helpman (1995), Krishna (1998, 2003). Levy (1997), McLaren (2002), Ornelas (2005a, b, c), Yi (1996), and Baldwin (1993).

13 Many adherents of the old view hold this premise so tightly that they insist on the using the term Preferential Trade Agreements (PTAs) in lieu of the WTO's official moniker, RTAs.

14 When the building-stumbling-block paradigm was established, unilateralism was not widely appreciated; the focus was on regionalism and its impact on the on-going multilateral negotiations (Uruguay Round). Subsequent academic literature that endogenised the reaction to exogenous regionalism typically employed political economy models where unilateralism never occurs, so unilateralism was excluded by modelling choices.For example, Reizman (1985), Kennan and Reizman (1990), Krishna (1998), Freund (2000a, b), Limão (2006), and Levy (1997). A later literature on "tariff complementarity" addresses this, but the results are routinely characterised as informing the regionalism-versus-multilateralism choice that is central the building-stumbling-block framing of the issues. See Bagwell and Staiger (1999a), Cadot et al. (1999), Freund (2000a), Yi (2000), Ornelas (2005a, 2005c), Bond et al. (2004) and Saggi and Yildiz (2009).

15 For an early and very explicit statement see Bhagwati (1993 p. 3); Panagariya (2000 p.312) provides a recent restatement. The seminal theory contributions – Limão (2006), Levy (1997), Freund (2000a), and Yi (2000) – embrace this exogenous endogenous distinction. Some writers have softened the exogenous/endogenous dividing line (e.g. Grossman and Helpman 1995, Krishna 1998). A later literature on "tariff complementarity" addresses this, but the results are routinely characterised as informing the regionalism-versus-multilateralism choice that is central the building-stumbling-block framing of the issues (instead of regionalism versus unilateralism). See Bagwell and Staiger (1999a), Cadot et al. (1999), Freund (2000a), Yi (2000), Ornelas (2005a, 2005c), Bond et al. (2004) and Saggi and Yildiz (2009).

different suppliers. There are only three elemental effects. First is "Smith's Certitude".[16] Nations that receive tariff preferences see a higher export price and thus export more to the preference giving nation. The preference-receiving nation (and especially its exporters) gains from this "trade creation". Second is "Haberler's Spillover".[17] Nations excluded from the preferences see lower export prices (they lower prices to remain competitive even though they still pay the tariff) and thus export less to the preference-granting nation. Excluded nations (and especially their exporters) lose from this "trade diversion". Third is "Viner's Ambiguity".[18] The preference-granting nation might or might not gain for the simple reason that preferences create a new distortion (tax discrimination among foreign suppliers) while removing another (tax discrimination between firms in the preference-granting and preference-receiving nations).

A critical point in interpreting the evidence flows from the first two effects. If trade flows are driven by preferential tariffs – or measures that are acting like preferential tariffs – we must observe trade creation driving trade diversion.[19]

This point is critical to interpreting the evidence, but not widely appreciated, so it is worth considering some non-preferential trade reforms that could lead to trade creation without trade diversion, or trade diversion that works in the 'wrong' direction. For example, suppose an RTA induces the integrating nations to apply identical safety standards for elevators, so that it is easier for the partners to sell to each other. As product regulations are general policies and not subject to rules of origin, it is logically possible that the RTA's regulatory standardisation makes it easier for third nations to sell to the combined RTA market. This could result in trade creation among the partners and "reverse trade diversion" for third nations (or what is sometimes called external trade creation).

The empirical evidence on trade diversion

As argued above, the Litmus Test for the relevance of Vinerian economics is the simultaneous observation of trade creation and trade diversion. If this is not observed, the provisions in the RTA are not acting like tariff preferences and Vinerian economics is inadequate.

A recent review, Freund and Ornelas (2010), note that most scholars estimate the increase of imports among RTA partners and reduction of imports from third nations using the gravity equation. This permits the researcher to control for other factors such as changes in income and various idiosyncratic factors related to the year, the bilateral relationship, and the individual nation. Summing up their review of the evidence, Freund and Ornelas write: "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." Nevertheless, most studies find evidence for trade creation. Taken together this suggests that it may be more correct to view RTAs as general trade liberalisations schemes than as discriminatory liberalisations.

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 EU, and 1992 bilateral RTAs between the EU and several Central European nations, Mercosur, ASEAN, and select RTAs among Latin American and among African nations. Only 8 of the 15 RTAs are found to have been effective in the sense of creating new trade among the partners (controlling for other factors).20 The others did not increase or decrease trade significantly. As trade creation is a prerequisite for preferential tariffs having an impact, these 8 RTAs with positive trade creation are the only ones where preferential tariffs might have been important. Of the 8 only 2 – the 1986 EU enlargement and

16 Smith (1776), as quoted in Pomfret (1997) puts it crisply: When a nation "exempt[s] the good of one country from duties to which it subjects those of all other … the merchants and manufacturers of the country whose commerce is so favoured must necessarily derive great advantage."

17 The discussion in Haberler (1936) runs over several pages but is centred on page 384. Haberler's spillover was certainly understood by scholars before Haberler (e.g. Bismarck used this aspect of customs union to force/cajole many German-speaking states to join his unified Germany), but I assign it to Haberler since Haberler's 1936 book shows that mainstream trade economists were confused about the theory of the second best, illustrating why Viner's 1950 book was viewed as such a landmark.

18 Viner, who framed his arguments in words alone (he was blissfully ignorant of postwar mathematical and diagrammatic analysis), couched his argument in the enduring but imprecise concepts of 'trade diversion' and 'trade creation'. Meade (1955) showed that the deep fundamentals turned on two effects: the trade volume effect (i.e. changes in trade over a domestic-versus-border price wedge), and the trade price effect (i.e. the terms of trade effect).

19 Specifically, the preference is why partner-nation exporters receive a higher price for their exports to the preference-granting nations. This producer price rise is why they export more, and these additional exports – i.e. trade creation – is what drives down domestic prices in the preference-granting nation. As third-nation exporters must match the new lower price inside the preference-granting nation, they see their producer price fall and this leads them to sell less (trade diversion). In short, trade creation drives trade diversion. Trade diversion without trade creation must be driven by something other than preferential tariffs or things that can be analysed as if they were preferential tariffs.

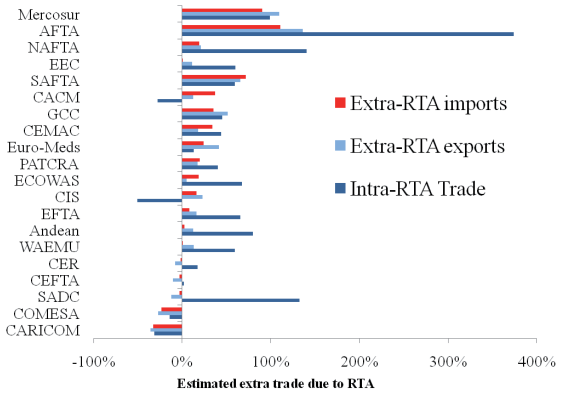
20 The 8 are the 1986 EU enlargement, the Andean Community, Mercosur, the ASEAN FTA, the EU's and EFTA's Association Agreements with Central and Eastern European countries, and NAFTA.

the EU Association agreements of the early 1990s are found to divert trade.

Moreover, Magee finds that NAFTA – the second largest RTA whose internal trade accounts for about 8% of world trade – created trade and produced "reverse trade diversion" – that is controlling for other factors NAFTA made it easier, not harder, for excluded nations to export to the US, Canada and Mexico. This result strongly suggests that using Vinerian economics to think about NAFTA's economic effects is misguided; NAFTA provisions do not seem to be acting like preferential tariffs in the aggregate.

Estimates based on 21st century data – where we have direct evidence on the small size of preference margins – show even weaker support for all-RTAs-are-PTAs view. A recent paper estimates that most RTAs have produced the opposite of trade diversion (Archarya, Crawford and Renard 2010).

Figure 10 Recent estimates of trade creation and trade diversion.



The estimated coefficients for trade creation (trade changes within the RTA), trade diversion (changes in imports from third nations), and new exports to third nations are plotted in Figure 10. The most striking finding is that almost all the RTAs have lead to external trade creation, i.e. reverse trade diversion – just the opposite of what would be expected if preferential tariffs were the main change imposed by the RTAs. Indeed, in the only RTAs where there has been large trade diversion, there has also been negative trade creation. Taken together, this suggests that RTAs are acting more like general trade liberalisation schemes, or in the case of COMESA and CARICOM as general trade restriction schemes.

These are critical pieces of evidence against viewing RTAs as being mostly or uniquely about tariff preferences. If RTAs were only about preferential tariffs, Magee should have found evidence of statistically significant trade diversion in all the RTAs where he also found trade creation. The world still has preferential tariffs, and trade diversion does occur in some sectors in some agreements, but the weight of empirical evidence now suggests that Vinerian economics is not sufficient for understanding the economic impact of RTAs.

The second pillar of the old approach also has serious problems when it comes to 21st century regionalism.

4.2 Regionalism vs multilateralism is inadequate

When the building-stumbling-block phrase was coined in 1991, the second pillar seemed a reasonable assumption, i.e. that the issue was regionalism-versus-multilateralism with the wave of regionalism taken as exogenous. After all, unilateralism was a rare bird in those days. Virtually all tariff cutting had been done in multilateral trade negotiations (MTNs) or in RTAs; reciprocity was critical to the politics of both both (see e.g. Destler 2005). Moreover, the avalanche of RTAs (see the light coloured bars in Figure 11) seemed to come out of the blue; taking the development as exogenous followed naturally.

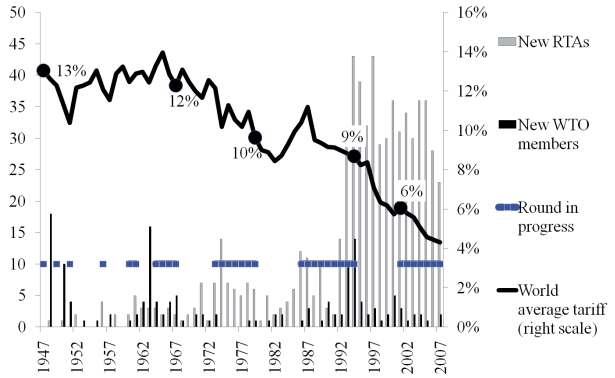
There are two major problems with the second pillar when it comes to 21st century regionalism. The most obvious is the prevalence of unilateralism (demonstrated in Section 3.4). To put it differently, if the observed regionalism since 1994 had not been accompanied by massive unilateral tariff cutting – i.e. the dark line in Figure 11 had been flat in the new century instead of falling – the old approach might have been useful. No MTN occurred while RTAs boomed so the old approach would have asserted that the RTAs were stumbling blocks. But the world tariff average did fall in the new century, due mostly to unilateralism.[21]

The second pillar could be restored if folding unilateralism into multilateralism was an option. But this too falls down on the facts. What the world has seen is an explosion of unilateral and regionalism, not unilateralism versus regionalism. Section 3 argued that regionalism and unilateralism seemed to be jointly endogenous reactions to a common third cause (the ICT triggered 2nd unbundling). The old framing of the question misdirects attention by highlighting a trade off that the world never faced.

21 About half of world trade is now covered by an RTA of some form, but the NAFTA was the last big one; after that, the RTA trade share was about 45% so the hundreds of RTAs expanded coverage by only 5% and many of these excluded most high-tariff items.

The second major problem for the second pillar is posed by a different set of facts – the historical parallelism of regionalism and multilateralism. For example, the US and Canadian stances on starting FTA negotiations were shaped primarily by the outcome of a political economy conflict inside each nation between national exporters (who would benefit) and national import competitors (who would lose). Their stance on an MTN was shaped by the same factors and actors. Given this, it is easy to understand the synchronicity of tariff-cutting decisions. US and Canadian exporters were in the ascendency in 1986 and they got their governments to embrace trade liberalisation regionally and multilaterally – the Canada-US FTA talks *and* the Uruguay Round were both started in 1986 (Schott and Smith, 1988).

Figure 11 Tariff liberalisation since 1947: RTAs, MTNs and unilateralism.



In fact, synchronicity between regional and multilateral tariff cutting has been the hallmark of the postwar period. In addition to the 1986, regional and multilateral tariff cutting marched in tandem in 1964. That was the year the Kennedy Round was launched and it was also the year the US and Canada started talks that led to the US's first postwar preferential trade liberalisation arrangement – the US Canada Auto Pact – which grew into one of the world's largest RTA outside of Europe (Keeley 1983 p.281). Two other GATT participants, Australia and New Zealand, also signed an RTA in that year. 1972 again witnessed parallelism. The EEC concluded talks for it first enlargement and bilateral FTAs with all non-acceding West Europeans in 1972 – the same year that the US, EEC, and Japan jointly called for a new GATT Round (Tokyo Round).[22] See Box 1 for details.

Since 1994, no new progress has been made on multilateral tariff cutting, but a different form of parallelism has continued. As Figure 11 shows, spreading regionalism has been associated with rapid unilateral tariff cutting (also see Section 3.4).[23] These facts tell us that that it is wrong to view the 1990s explosion of regionalism as exogenous.

This more fundamental analysis suggests that the old building-stumbling-blocks approach is logically misstated – it asks about the correlation of two endogenous variables driven by common factors (illustrated by the double-headed arrow in the diagram).

22 On the time of the Tokyo Round agreement, see Low (1993 p. 176).

23 The unilateralism can be seen as the continued fall of the world tariff average despite the lack of MTN cutting and despite the fact that regionalism since 1995 has involved only small shares of world trade; see Section 3 on these facts.

Box 1 Parallelism in RTAs and GATT Rounds

From the GATT's formation right up to the end of the 20th century, regional and multilateral liberalisations have displayed a remarkable parallelism.

- 1947 saw the creation of the GATT and the Organisation for European Economic Cooperation (OEEC) that coordinated preferential European liberalisation up to the end of the 1950s (particularly of intra-European bilateral quantitative restrictions).
- Annecy GATT Round was opened in April 1949 – the same year that European preferential integration took off with the OEEC's European Payments Union.
- The 1950-51 Torquay Round was launched just a few months after the famous Schuman Declaration that is viewed by many as the birthday of today's European Union.
- When European regional tariff cutting got serious from 1960, GATT tariff cutting restarted first with the Dillion Round (1960-1961).
- In 1964 – the year the Kennedy Round was launched – the US and Canada started talks that led to the US's first postwar preferential trade liberalisation arrangement – the US Canada Auto Pact – which grew into one of the world's largest PTAs (Keeley 1983 p.281). Two other GATT participants, Australia and New Zealand, signed their own PTA.
- 1972 again witnessed parallelism between regionalism and multilateralism. The EEC concluded talks for it first enlargement and bilateral FTAs with all non-acceding West Europeans. The same year, the US, EEC, and Japan jointly called for a new Round, the Tokyo Round (Low 1993 p. 176).
- 1986 saw two big steps in preferential liberalisation and a new Round. The US and Canada launched talks to broaden their 1965 sectoral arrangement into a deep FTA while Europe launched its massive deepening preferential trade (Single Market programme). The Uruguay Round was also launched in 1986 with an agenda that included many 'deep' issues.

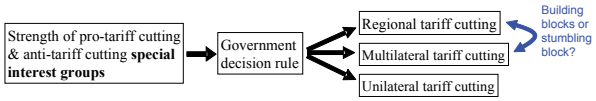
The parallelism broke down in the late 1990s. After some early post-Uruguay Round successes in 1997 – e.g. the Information Technology Agreement that created global free trade in many ITC goods – multilateral liberalisation stagnated while preferential trade liberalisation went from strength to strength. The many years it took to launch the Doha Round and the decade of negotiations to date gave preferentialism a very clear lead in the minds of many WTO members.

4.2.1 The old approach's logical incompleteness

Delving more deeply into the old approach's problems, it is clear that it is logically incomplete. Surely government decisions on RTAs are also endogenous to political economy forces, so it is logically incomplete to take regionalism as exogenous. A more complete analysis is suggested by the endogenous trade policy literature (e.g. Grossman and Helpman 2002). This approach is premised on the assertion that special interest groups are the fundamental determinants of trade policy. This approach has quite different implications for how one thinks about regionalism and multilateralism.

The basic point can be illustrated with Figure 12. The three forms of tariff cutting are shown at the right-side, namely regionalism, multilateralism, and unilateralism. The determinants of all three forms of tariff reform are shown with the straight black arrows.

Figure 12 Endogenous-trade-policy approach to the tariff cutting..



Special interest politics shape the government's decision rule which in turn explains tariff reform choices and the form they take (RTA, MTN or unilateral). The ultimate determinant of a nation's stance on regionalism, multilateralism and unilateral is the strength of pro- and anti-liberalisation groups.

Most writers in this tradition focus on the economic power of exporters and import-competitors and how this power filters through to governmental decision rules via the national political system.[24] In this sense, the economic power of export firms and import-competing firms are the ultimate causes of trade policy choices. This applies equally to regional, multilateral, and unilateral tariff cutting.

5. What would the new approach look like?

Up to the late 1980s regionalism was limited. Economists such as Jacob Viner, James Meade, and Harry Johnson focused on narrow questions, chiefly: "Would a particular nation gain from joining a free trade agreement?" New facts appeared in the 1990s and new thinking was needed. The building-stumbling-block approach was developed to think about shallow RTAs and MTNs.

The 21st century has thrown up new facts and now a new framework for thinking about them is needed. As the previous section argued, the 1990s approach is an inadequate framework for thinking about 21st century regionalism and its implications for the world trade system. To summarise, its focus on tariff preferences is not appropriate for 21st century regionalism. Its focus on the regionalism-versus-multilateralism distinction is also amiss in a world where unilateralism is a key driver of trade opening. Finally, the old approach was always logically incomplete in that focused on correlations between endogenous variables (RTA and MTN outcomes) rather than on the political economy roots of trade liberalisation.

Arguing that the old building-stumbling-block logic is not up to the job of understanding 21st century regionalism is much easier than finding a replacement. What the world needs is a new framework that is as simple and compelling as the old one, but relevant to 21st century regionalism. Such a framework has not yet emerged. It is useful, nevertheless, to list the features it should include. The starting point is a discussion of the key features of deep RTA provisions.

5.1 Economics of deep RTA provisions

Much of 21st century regionalism concerns regulatory measures, not tariff measures. This means that the fulcrum of the new thinking will be regulatory economics, not Vinerian economics. Unfortunately, regulatory economics is harder and much less flexible than Vinerian economics since it often has to deal with industry-specific characteristics. For example, in telecoms, network externalities are a dominate concern, but less so in the air cargo. Asymmetric information is a prime concern in financial services but less so in capital movement regulations. In some sectors, imperfect competition considerations are central while in other sectors they are not. And the list goes on. This suggests that it may not be possible even for the most brilliant economist to lay out the basic analytics as cleanly as Paul Krugman did for the old thinking in his 1991 articles.[25]

5.1.1 Hints from the lack of trade diversion

One critical fact that the new framework must explain is the presence of "reverse" trade diversion, i.e. that RTAs seem to be making it easier for imports from partner and third nations. As it turns out, many of the deep RTA provisions

24 An early reference is Cooper (1971 p.410), but also see Roesseler (1978), Blackhurst (1979), Baldwin (1980), Moser (1990), or Hillman and Moser (1992). The best known reference is the trade-wars-trade-talks paper by Grossman and Helpman (1995).

25 Krugman (1991a,b)

are not naturally thought of as preferential despite their being reforms struck in the context of an RTA. Before turning to specific deeper RTA provisions, consider the general source of the non-discrimination.

General considerations: Nationality of firms and pubic good features

As it turns out, it is hard, or even impossible, to write rules of origin for many deep provisions. The trouble-making factors concern: 1) the difficulties of determining the nationality of 21st century companies, 2) the public good nature of infrastructure services, and 3) the public-good nature of regulatory reform.

The first point is easy. In today's world, it is difficult to establish a company's nationality so it is difficult to write deep RTA provisions that only apply to companies of any particular nation. As we shall see in the examples below, this problem has frequently led RTAs to define the affected firms by where they are incorporated. This allows firms from third nations to free ride on bilateral opening by incorporating affiliates in one of the RTA nations.

The second point is more subtle. As argued in Section 3, deep RTAs typically foster the trade-investment-services nexus by liberalising developing countries' infrastructure service sectors, for example, telecoms. The RTA may be written in a way that gives telecom companies from one partner an edge in providing telecom services in the market of the other. In this sense it is discriminatory and may divert services trade from, say, a US telecom to, say, a Japanese telecom. But as far as trade in goods is concerned, it matters little whether the world-class telecoms are provided by a US or Japanese company. Trading firms from all nations find it easier to sell to the developing country member of the RTA when telecommunications work well. In this way, a deep RTA provision on telecom liberalisation can act like a public good for exporters from all nations.

The third point is related. One important set of BBBs concerns health, safety and environmental regulations. Given these are typically highly technical, governments frequently ask the regulated industry to write the regulations (or at least gives them a big say in their formulation). This naturally results in rules that favour incumbent domestic firms at the expense of all others. To make the nation more attractive to foreign investment, the nation may reform its regulations in the context of an RTA. For example, the RTA may commit the nation to rely on international standards instead of national standards. The political economy logic of this provision may be bilateral, but its effect can be multilateral. Disciplining idiosyncratic regulations is likely to make it easier for exporters from all nations to sell to RTA markets.

Discrimination possibilities in six types of deep RTA provisions

To flesh out these general points, consider the relevance of Vinerian economics to the analysis of six major areas of deeper RTA provisions: trade in services, government procurement, competition policy, investment performance measures, technical barriers to trade, and trade remedies.[26]

Barriers to trade in services rarely occur at the border. Establishing discrimination – the sine qua non of the applicability of Vinerian economics – is particularly difficult for two reasons. First, it is hard to establish the origin of a service. Just think of how hard it would be to trace the nationality of all the value added in a financial service or a telecom service.[27] In reaction, government may turn from determining the origin of the service to determining the origin of the service provider. This tactic, however, runs into the second intractable problem. Corporate nationality is a hard thing to define precisely – especially for the large firms that dominate global trade. In reaction, governments often turn to legalistic definitions such as the nation where the company is incorporated. However once the rules-of-origin have reached this point, they are easy to get around. For example, if the rule of origin in the Japan-Malaysia EPA provides access only to banks registered in Japan or Malaysia, the rule encompasses all the US and European banks that have affiliates incorporated in Tokyo.

Government procurement provisions are easy to subject to rules of origin as long as they concern traded goods. For these provisions, Vinerian economics works fine. But much of government procurement applies to services and here to rule-of-origin issues mentioned above come to the fore.

Competition policy provisions are typically aimed at assuring partner companies that their market access and/or investments will not be threatened by anti-competitive practices of private or state-owned local firms. Competition policy, however, is difficult to subject to rules of origin. Part of the problem is that it is difficult to establish the nationality of modern corporations, as mentioned above. Another part stems from regulatory nature of most competition policy. In most nations, the competition policy is based on rules delimiting forbidden behaviours (e.g.

26 See Baldwin, Evenett and Low (2009) for a more detailed analysis of specific deep RTA provisions.

27 This is a well-known problem and the same issue bedevils VAT authorities when it comes to de-taxing service exports.

price fixing) without regard to the nationality of the defendant or plaintiff. Thus is Korea's competition policy is strengthened by the EU-Korea FTA, firms from third nations benefit as well from the resulting increase in fairness of the Korean market place.

Investment performance provisions typically assure foreign investors that the host government will not interfere with business decisions. It is difficult to make this discriminatory for the reason mentioned above: difficulties in defining corporate nationality, Moreover, governments often embrace such provisions as part of general move to a more business-friendly investment climate and so end the practice for all foreign firms, not just those coming from RTA partners. Indeed, developing nations may make these explicitly nationality-neutral, so as to diffuse their reliance on FDI from their major source.

Technical barriers to trade (TBT) – such and health, safety, and environmental regulations and standards – are also difficult to apply on a discriminatory basis, as the EU's experience with its radical TBT liberalisation (the Single Market Programme) shows. TBTs are usually justified on the basis of 'good governance' criteria, like protecting consumer's health. Such goals, however, admit little room for discriminating product by origin – either a product is safe or it is not, regardless of where it is made.

Finally, RTA provisions concerning anti-dumping and anti-subsidy tariffs concern trade in goods and so discrimination is possible using standard rules of origin. Here Vinerian economics is appropriate.

To summarise, many but not all deeper provisions tend to act as general liberalisations rather than discriminatory liberalisations because it is difficult or impossible to write rules of origin for them that exclude third nations. The deep reasons are the difficulties in establishing the nationality of modern corporations and of services as well as the public-good nature of the many regulatory reforms in deep RTAs.

5.1.1 Political economy of 21st century RTAs

The new framework will also have to suggest a new political economy. The standard political economy logic applied to trade liberalisation deals with tariffs and it focuses on an exchange of market access.[28] This seems inadequate when explaining agreements whose main goal is to establish disciplines that foster the trade-investment-services nexus. As discussed in Section 3.1, the nexus requires firms to connect factories and do business abroad. Deep RTAs provide assurances on both fronts. The bargain in a 21st century RTA is "foreign factories in exchange for domestic reforms", not, "exchange of market access" as was the case for 20th century RTAs. Of course, market access is still important, but the deep provisions are not really about market access – they are about helping foreign companies connect production facilities internationally, and do business locally.

Likewise, the winners and losers from 21st century RTAs invoke a much richer set of players, so the new political economy framework will have to fit in a broader range of political economy actors. For example, in many cases the losers from reform may be incumbent corporations while the winners will be all other corporations, domestic and foreign.

An important implication of this recasting of the basic political deal is that not all nations can drive such a bargain. Only nations that possess technology that they are willing to offshore have the leverage to demand the massive domestic reforms that are so common in 21st RTAs. To date, that basically boils down to US, the EU and Japan although Korea and Taipei, China may also fall in this category. South-South FDI flows are rising, and South-South BITs along with them, so this feature many change.

The new political economy framework must also account for all three forms of trade liberalisation – multilateralism, regionalism, and unilateralism. In particular, the landmark fact has to be the parallelism of RTAs, MTNs and unilateralism.

5.1.2 Parallelism as explained by the juggernaut effect

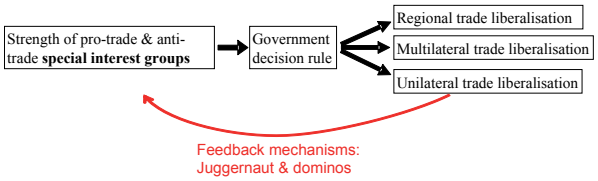
The parallelism of RTAs and MTNs up to the mid-1990s and the fact that the parallelism continued after 1995 in the form of RTA and unilateralism provide important hints as to how we should think about global trade liberalisation in general and 21st century regionalism in particular. Indeed, the two forms of parallelism are clear indications that something is altering the fundamental determinants of trade policy in a cyclic fashion. One line of thinking points the finger to earlier trade liberalisation as the cause of the cyclic behaviour. This is illustrated in Figure 13.

The basic idea is that tariff reforms themselves trigger feedback mechanisms that alter the state of pro- and anti-liberalisation forces in each nation. Specifically, reciprocal trade liberalisation strengthens pro-trade political economy forces and weakens those of anti-trade forces.[29] As tariffs

28 See, for example, Grossman and Helpman (1995), or Baldwin and Baldwin (1996).

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

Figure 13 Juggernaut and parallelism in RTAs, MTNs and unilateralism



come down reciprocally – either multilaterally or regionally – import competing sectors get smaller and typically less influential in trade policy formulation. Similarly, as trade partners lower their tariffs, exporters grow in size and political strength with the improved access to foreign markets. In short, this 'juggernaut' feedback mechanism asserts that reciprocal liberalisation tends to reshape the political economy landscape inside each nation in a way that makes future liberalisation more likely.

The reciprocal liberalisation is typically phased in over 5 to 10 years and the necessary entry (of exporters) and exit (of import-competitors) may take even longer. Under the juggernaut approach, this is the source of the time dimension which leads to the episodic, synchronised nature of trade liberalisation according to the juggernaut logic. Once the tariff-cutting ball starts rolling, political economy momentum keeps it rolling until all tariffs in its path are crushed.[30]

A second feedback mechanism is the so-called domino effect whereby the signing of one RTA tends to induce the signing of more RTAs. The basic notion is that as FTAs reduce the exports of third nations to the integrating nations, they also stimulate third-nation exporters to engage in new political economy efforts to get their government to redress the new discrimination. In many cases, third-nation governments respond by signing new FTAs with one or both of the partners, who have recently integrated.31

5.2 Political economy of unilateralism and the second unbundling

The new approach must give a central place to unilateralism as it is one of the three mainstays of 21st century regionalism. Economists' thinking on unilateralism is not highly developed, but a number of explanations have been offered, some of them focusing squarely on the trade-investment-service nexus.[32] These explanations focus on the suggestive fact that unilateralism boomed at approximately the same time as supply chains started to stretch across North-South borders.

One story is that emerging markets cut tariffs on parts and components to attract foreign factories and got caught in a "race to the bottom" (Vézina 2010). Another turns on the logic of effective rates of protection. Cutting tariffs on an imported input raises the level of effective protection on downstream products (Cordon 1966). The political economy of this explains why most nations have lower tariffs on imported intermediates than final goods – this sort of tariff cascading is a way of increasing protection of the downstream activity (see Cadot, de Melo and Olarreaga 2004). As the 2nd unbundling proceeded, almost everything became a traded intermediate input and so faced new pressures for liberalisation. While there may be some merit in these ideas, it is clear that there is much work to be done on the political economy and empirics of unilateralism.

5.2.1 Fiscal federalism and the WTO "border question"

When thinking about tariffs, it is clear that multilateralism is most efficient. But when it comes to the deeper disciplines addressed by 21st century regionalism, the most efficient level of governance is less clear. The theory of fiscal federalism should be part of the new framework's thinking when it comes to the optimal allocation of policy setting.

There is a good analogy with what goes on in the EU where the question is: What policies should be decided at the national level and which at the supranational level? The answer in the EU is that the appropriate level depends upon the nature of the discipline. When it comes to food standards, general guidelines are agreed at the EU level, but detailed regulations are decided at the national level. Tariff and competition policy, however, are allocated to the EU level.

Similar thinking should be developed to make recommendations on which of the disciplines that underpin the trade-investment-service

30 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 once in motion, it is hard to stop. 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.

31 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; empirical support is provided by Egger and Larch (2008) and Baldwin and Jaimovich (2009). On the theory, also see Bond and Syropoulos (1996), Freund (2000), Yi (1996), McLaren (2002), Levy (1997), and Krishna (1998).

32 In the economics literature, most discussions of unilateralism consist of practical accounts of how and why various nations undertook such measures (e.g. Garnaut 1991, Young 1996, Edwards and Lederman 1998, Richardson 2001, and Sally 2008). The political economy theories that account for unilateralism include Coates and Ludema (2001), Krishna and Mitra (2008), Baldwin (2005b), and very recently Ludema, Mayda and Mishra (2010), Conconi and Perroni (2010), and Baldwin (2010). This section draws on the latter.

nexus should be agreed at the multilateral level, at the regional level, and at the national level.

6. Implications for the world trade system

21st century regionalism is good news and bad news for the world trade system. The good-news part is easy to explain. Trade liberalisation has progressed with historically unprecedented speed in the 21st century, even when measured by a 20th-century-trade yardstick like average tariffs (Figure 11). As a result, trade volumes have boomed, lifting billions out of dire poverty. Twenty years ago, one could wonder whether regionalism would be a building or stumbling block; now we know there were no stumbling blocks on the road to zero tariffs. The road remained open and the world is driving down it as fast as ever. This building-stumbling-block thinking, however, focuses attention on the wrong issues.

21st century regionalism is a threat to the world trade system, but the nature of the threat is subtle.

6.1 Regionalism's real threat to global trade system

21st century regionalism has three parts: Deep RTAs, BITs, and unilateralism. Unilateralism is not a systemic threat to the WTO – although it does make finishing Doha harder. Likewise, the BITs have co-existed with the WTO for decades without any apparent harmful spillovers.

The real threat is that deep RTAs may undermine the WTO's central place in world trade governance. But the threat to WTO centricity does not come on the tariff-cutting front; it comes on the rules-writing front. As was shown above, the tariffs have come down a great deal already, although they are still important in some sectors, especially agriculture.

More specifically, deep RTAs may undermine the WTO as the forum for agreeing new rules – specifically the rules necessary to foster the trade-investment-services nexus that is the core of today's international commerce. But why is the choice of forum a problem for global trade governance? There are three main reasons to worry about the WTO being sidelined on the rule writing front.

First, the basic WTO trade norms are almost universally accepted and respected – a very rare thing (think of climate change, nuclear proliferation, or human rights). These norms are a global public good of enormous, if unquantifiable benefit. The universality of the norms stems in large part from the way they were promulgated – in multilateral negotiations where the GATT/WTO consensus principle held sway. The new trade disciplines are being promulgated in settings of massive power asymmetries – the deep RTAs signed by the US, EU and Japan with small to medium sized developing nations. Lacking the legitimacy that comes from multilateralism and consensus, it is not at all clear that the new norms will be universally respected.

For example, some emerging markets – China, India and Brazil – are large enough to attract foreign investment and technology without signing deep RTAs, and they have so far shunned them.[33] China in particular might decide to reject the rules – creating something like a "Cold War of deeper trade disciplines". This sort of distrust could spread beyond the new rules, especially if China, India and Brazil feel that the US is practicing 'competitive liberalisation' – trying to encircle them in a way that eventually confronts them with what might be seen as an ultimatum. This outcome would be made more likely if the US reverts to its aggressive unilateralism of the 1980s (the Plaza Accord and the Structural Impediments Initiative that forced Japan to revalue and remove behind-the-border barriers) and 1970s (Nixon 10% surcharge that forced Germany to revalue).

Second, a world where the WTO's importance starts to resemble that of UNCTAD – with all the action going on in RTAs – is not a world that fosters multilateral cooperation on other issues, such as trade-related policies that help with climate mitigation and adaption, or food shortages linked to drought or floods. US, EU and Japanese interests will be served in the short term, and the interests of small to medium emerging markets will likewise be served (if not evenly), but where do Brazil, India and China fit in?

These nations are not in a position to set up their own systems of deeper disciplines for the trade-investment-services nexus because they do not have advanced technology factories to offshore in exchange for host-nation reforms (on the political economy of 21st century regionalism, see Section 5.2). By the time their multinationals are ready to make major outward pushes, the rules-of-the-road written by the US, EU and Japan will have been firmly embedded into international commerce. More precisely, they will be embedded in the domestic laws and regulations of all the host-nations that the Chinese, Indian and Brazilian companies will be looking at. Like it or not, Chinese, Indian and Brazilian companies will have to play by the

33 The EU-India agreement, for example, excludes many of the deeper disciplines in the EU's other RTAs.

rules that are now being written by the US, EU, and Japan in agreements that involve massive power asymmetries.

If Brazil, India and China play their assigned roles in this storyline, it may all work out peacefully. But that is not the only outcome observed when such tactics were applied historically. This is a world that starts to resemble the 19th Great Powers situation. That episode of globalisation did not end well.

This is not the only scenario, of course. A whole system of trade and investment disciplines has developed in the form of the BITs. Up to now, the BITs and its system of jurisprudence, negotiations and politics does not seem to have undermined the WTO's authority on the issues covered in the 1995 Marrakesh Agreement. But as international commerce becomes ever more dominated by the trade-investment-services nexus, the WTO may be increasingly sidelined when it comes to trade governance.

Third, the WTO's adjudication function is still working well, but any dispute settlement system must walk on two legs. The judges can connect the dots for particular cases, but the basic rules must be updated occasionally to match evolving realities. For example, the Appellate Body finds itself ruling on issues like "zeroing" where the negotiated consensus is disputed. If the basic rules applied by the Appellate Body are not updated, there is a serious danger that the judges will overreach themselves, basing decisions on previous decisions that were based on previous decisions. Similar challenges may arise when members ask the Appellate Body to rule on 21st century climate subsidies and taxes based on rules negotiated in the 1940s and last updated in 1994. The larger members may be tempted to take matters into their own hands, applying sanctions based on unilateral law, not multilateral law.

Finally, if the WTO gets sufficiently sideline, it may prove very difficult to successfully negotiate rule updates. Hereto, the GATT/WTO has always packaged rule-updating in Rounds that were driven primarily by trade-liberalisation politics (juggernaut). If all the trade liberalising action moves to the RTAs, WTO members will have to find a new way to negotiate rule updates.

6.2 What is to be done?

One course of preventive action would be to work towards multilateralising the deeper disciplines in RTAs. Many of the deep disciplines (e.g. opening up of the telecom sector, or liberalisation of the air transport sector) are embodied in national laws. Many of these have a public-good nature to them in the sense that they facilitate trade with all nations, not just the members of the RTA which brought about the reform. Questions of consistency seem to be of second order for such measures, but for other measures, e.g. intellectual property protection, or investors' rights, it is not clear that the various deeper disciplines are compatible.

Distinguishing the various categories of disciplines is an important task for trade scholars and governments. The WTO's centricity is not in peril if the various deep RTAs turn out to have implemented reforms that are consistent with each other. Such disciplines, or at least the basics, might be multilateralised with WTO agreements (like the GATS), or plurilateral agreements like 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 think about 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.

One hint in this direction is the near universality of certain provisions. Data from the WTO's World Trade Report suggests that there are four core disciplines in deep RTAs that go beyond WTO agreements. These are competition policy (covered by 47% of all agreements in the WTO database), movement of capital (39%), intellectual property rights not in the TRIPs Agreement (37%), and investment not covered by GATT 1994 (31%).

More modest versions of multilateralising the deeper disciplines 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 – e.g. FTAs should provide substantial sectoral coverage, substantially eliminate discrimination in national treatment in the affected sectors, and raise no barriers against third nations. Even these very basic guidelines are completely absent when it comes to deeper provisions like competition policy, rights of establishment, FDI-linked capital flows, etc. Since many of the deep RTAs are aimed at improving the investment climate rather than providing discriminatory market access, many of the deep RTA provisions already respect rules like

those for Services FTAs. Perhaps then it might not be too difficult to codify a set of guidelines in a WTO agreement, or plurilateral.

6.3 Concluding remarks

The rise of 21st century regionalism is not yet a disaster for the world trade system. It has kept trade liberalisation and trade booming despite the WTO's slow progress. But the present course of events seems certain to undermine the WTO's centricity – RTAs will take over as the main loci of global trade governance. Over the past ten years, WTO members have "voted with their feet" for the RTA option. Without a reform that brings existing RTA disciplines under the WTO's aegis and makes it easier to develop new disciplines inside the WTO system, the RTA trend will continue, further eroding WTO centricity and possibly taking it beyond the tipping point where nations ignore WTO rules since everyone else does.[34]

This scenario runs the risk that global trade governance drifts back towards a 19th century Great Powers world. In the best of cases, the WTO would continue to thrive as the institution that underpins 20th century trade flows. The Marrakesh agreements would form a 'first pillar' of a multi-pillar trade governance system. All the new issues would be addressed outside the WTO in a setting where power asymmetries are far less constrained. This is what has happened with the BITs – they established a parallel system of disciplines without substantially undermining the WTO's authority on Marrakesh disciplines. But this is not the only scenario. It is also possible that the WTO's inability to update its rules gradually undermines the authority of the Dispute Settlement Mechanism.

If the RTAs and their power asymmetries take over, there is a risk that the GATT/WTO would go down in future history books as a 70-year experiment where world trade was rules-based instead of power-based. It would, at least for a few more years, be a world where the world's rich nations write the new rules-of-the-road in settings marked by vast power asymmetries. This trend should worry all world leaders. In the first half of the 19th century, attempts by incumbent Great Powers to impose rules on emerging powers smoothed the path to humanity's greatest follies – the two world wars.

References

Aghion, Philippe, Pol Antras, Elhanan Helpman (2007)."Negotiating free trade," *Journal of International Economics*, Elsevier, vol. 73(1), pages 1-30, September.

Amador, João and Sónia Cabral (2008). "Vertical specialization across the world: A relative measure", Bank of Portugal Working paper, No. 10-2008.

Ando, Mitsuyo and Fukunari 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, Volume 14, pages 177-216.

Ariu, Andrea and Giordano Mion (2010). "Trade in services: IT and task content," Research series 201010-200, National Bank of Belgium.

Bagwell K, Staiger RW. (1999a). Regionalism and multilateral tariff cooperation. In *International Trade Policy and the Pacific Rim*, ed. J Piggott, A Woodland, pp. 157-185. New York: St. Martin's Press.

Balboa, Jenny (2008). "Negotiated Trade Liberalization in East Asia: Examining Japan's Economic Partnership Agreement (EPA): Focusing on the Japan Philippines Economic Partnership Agreement (JPEPA)", Philippine APEC Study Center Network, Philippines Institute for Development Studies working paper.

Baldwin, Richard (1993). 'A domino theory of regionalism', NBER WP, 4465, published in Baldwin, Haaparanta and Kiander (eds.), *Expanding Membership in the European Union*, Cambridge University Press.

Baldwin, Richard (1994). *Towards and integrated Europe*, CEPR, London.

Baldwin, Richard (2002). "Asian Regionalism: Promises and Pitfalls," in *East Asian Economic Regionalism: Feasibilities and Challenges*, C. Ahn, R.Baldwin, and I. Cheong (eds.), Springer, April 2005.

Baldwin, Richard (2004). "The spoke trap," CNAEC Research Series 04-02, KIEP, Seoul.

Baldwin, Richard (2006a). "Globalisation: the great unbundling(s)," Chapter 1, in Globalisation challenges for Europe, Secretariat of the Economic Council, Finnish Prime Minister's Office, Helsinki, 2006, pp 5-47.

Baldwin, Richard (2006b). "Managing the Noodle Bowl: The fragility of East Asian Regionalism," CEPR DP 5561, eventually published in

34 See more detailed arguments on this point see Baldwin (2008) and Baldwin and Carpenter (2009).

Singapore Economic Review, vol. 53, issue 03, pages 449-478, April 2008.

Baldwin, Richard (2008). "The WTO tipping point", VoxEU.org, 1 July 2008; http://www.voxeu.org/index.php?q=node/1345

Baldwin, Richard (2010) "Understanding the GATT's wins and WTO's woes", *CEPR Policy Insight* No.49, 5 June.

Baldwin, Richard and Robert Baldwin (1996). "Alternate approaches to the political economy of endogenous trade liberalization", with Robert E. Baldwin, *European Economic Review*, 40, pp 775-782.

Baldwin, Richard, Philippe Martin (1999). "Two waves of globalisation: Superficial similarities and fundamental differences", in H.Siebert (ed.), *Globalisation and Labour*, J.C.B. Mohr for Kiel Institute, Tubingen, 1999. Also NBER WP 6904.

Baldwin, Robert (1970). *Non-Tariff Distortions of International Trade*. Washington, D.C.: The Brookings Institution.

Baldwin, Robert (1980). "The Economics of the GATT" in *Issues in International Economics*, edited by Peter Oppenheimer. Stocksfield, England and Boston: Oriel.

Bergsten (1995). C. Fred Bergsten. "Competitive Liberalization and Global Free Trade: A Vision for the Early 21st Century." Institute for International Economics, working paper 96-15. Washington, D.C.

Bergsten, Fred (1996). "Globalizing Free Trade: A Vision for the Early 21st Century," *Foreign Affairs*, Vol. 71, No. 3, pp.105–20.

Bhagwati, Jagdish (1991). *The World Trading System at Risk*, Harvester Wheatsheaf, Hemel Hempstead.

Bhagwati, Jagdish (1993). "Regionalism and multilateralism: an overview," in J. de Melo and A. Panagariya (eds), *New Dimensions in Regional Integration*, Cambridge University Press, Cambridge.

Bond, Eric W., and Constantinos Syropoulos (1996a). "The size of trading blocs: Market power and world welfare effects," *Journal of International Economics*, Vol. 40, pp. 411–437.

Bond, Eric W., and Constantinos Syropoulos (1996b). "Trading Blocs and the Sustainability of Inter-regional Cooperation," forthcoming in Canzoneri, M., Ethier, W. and Grilli V. (eds.) *The New Transatlantic Economy*, Cambridge University Press, Cambridge.

Bond, Eric W., and Constantinos Syropoulos and Alan Winters (1996). "Deepening of Regional Integration and External Trade Relations," CEPR Discussion Paper No. 1317.

Cadot, Olivier & Jaime de Melo & Marcelo Olarreaga, 2004. "Lobbying, Counterlobbying, and the Structure of Tariff Protection in Poor and Rich Countries," *World Bank Economic Review*, Oxford University Press, vol. 18(3), pages 345-366.

Carpenter, Theresa and Andreas Lendle (2010). "How Preferential is World Trade?" HEID Working Paper 2010-32.

Coates, D., and Ludema, R., 2001, "A Theory of Trade Policy Leadership", *Journal of Development Economics* 65(1).

Conconi, Paola and Carlo Perroni (2010). "Conditional versus Unconditional Trade Concessions for Developing Countries", pdf dated September.

Cooper, Richard (1971). "Tariff issues and the Third World", *The World Today*, Chatham House publication, September.

Corden, W. M. (1966). "The Effective Protective Rate, the Uniform Tariff Equivalent, and the Average Tariff," *Economic Record* 42, 200-216.

Dallas Fed (2002). "Maquiladora Industry: Past, Present and Future." Federal Reserve Bank of Dallas, El Paso Branch, Issue 2.

Destler, I. M. (2005). *American Trade Politics*, 4th edition, Peterson Institute, June.

Egger, Peter, and V. Merlo. "The Impact of Bilateral Investment Treaties on FDI Dynamics." *The World Economy*, 30, 2007, 1536–1549.

Either, Willam (1996),'Multilateral roads to regionalism,' CIES DP, (University of Adelaide).

Ethier, William (2002). "Unilateralism in a Multilateral World," *Economic Journal*, Royal Economic Society, vol. 112(479), pages 266-292.

Feenstra, Robert and Gordon Hanson (1997). "Foreign direct investment and relative wages: Evidence from Mexico's maquiladoras," Journal of International Economics, Elsevier, vol. 42(3-4), pages 371-393, May.

Freund, Caroline & Ornelas, Emanuel, 2010. "Regional trade agreements," Policy Research Working Paper Series 5314, The World Bank.

Freund, Caroline (2000a) "Different Paths to Free Trade: The Gains from Regionalism". *Quarterly Journal of Economics* 115:4, 1317-1341.

Freund, Caroline (2000b). "Multilateralism and the endogenous formation of preferential trade agreements", *Journal of International Economics* 52 (2000) 359–376.

Fugazza, Marco and Frédéric Robert-Nicoud (2010). "The Emulator Effect of the Uruguay Round on US Regionalism," CEP Discussion Papers dp0973.

Garnaut, Ross (1991). Australia: A case study of unilateral trade liberalisation, in J. Bhagwati and H. Patrick, *Aggressive Unilateralism: America's 301 Trade Policy and the World Trading System*, University of Michigan Press.

Gaulier, Guillaume, Francoise Lemoine and Deniz Unal-Kesenci (2007). "China's emergence and the reorganisation of trade flows in Asia," *China Economic Review*, vol. 18(3), pages 209-243.

Grossman, G. and E. Helpman (1995). "The Politics of Free-Trade Agreements," *American Economic Review* 85 (1995b):667-690.

Grossman, Gene and Elhanan Helpman (2002). *Interest Groups and Trade Policy* , Princeton and Oxford UK: Princeton University Press.

Grubel, Herbert G.; Lloyd, Peter J. (1971). "The Empirical Measurement of Intra-Industry Trade". *Economic Record* 47 (4): 494–517.

Haberler, Gottfried (1936). The theory of international trade with its applications to commercial policy, W. Dodge & Company, New York.

Hillman, Arye and Peter Moser (1992). 'Trade Liberalization as Politically Optimal Exchange of Market Access' published in 1996 in *The new transatlantic economy,* Matthew B. Canzoneri, Wilfred Ethier, Vittorio Grilli (eds), Cambridge University Press.

Horn, Henrik, Petros Mavroidis, and André Sapir (2010). "Beyond the WTO? An Anatomy of EU and US Preferential Trade Agreements," *The World Economy*, Blackwell Publishing, vol. 33(11), pages 1565-1588

Hufbauer, G and J. Schott (1993). *NAFTA: An Assessment,* Revised Edition, Peterson Institute, October.

Hufbauer, G and J. Schott (2005). *NAFTA Revisited: Achievements and Challenges*, Peterson Institute, October.

Hufbauer, Gary and Matthew Adler (2011). Policy liberalisation and US integration with the global economy: trade and investment between 1980 and 2006, Chapter 17 in M. Jovanovic (ed) *International Handbook on Economic Integration*, Volume III, pp 371 – 416.

Ishii, Jun and Kei-Mu Yi (1997). "The growth of world trade," Research Paper 9718, Federal Reserve Bank of New York.

Johnson, Robert and Guillermo Noguera (2010). "Accounting for Intermediates: Production Sharing and Trade in Value Added", manuscript.

Kimura (2005)

Kimura, Fukunari, Yuya Takahashi, and Kazunobu Hayakawa (2007). "Fragmentation and parts and components trade: Comparison between East Asia and Europe", The North American *Journal of Economics and Finance*, Volume 18, Issue 1, 1 Pages 23-40.

Krishna Pravin (2003). Are regional trading partners "natural"? *Journal of Political Economy* 111:202-26

Krishna, Pravin (1998). "Regionalism and Multilateralism: A Political Economy Approach," *Quarterly Journal of Economics*, February 1998, Vol. 113, No. 1, Pages 227-250.

Krishna, Pravin and Mitra, Devashish (2008). "Reciprocated unilateralism in trade reforms with majority voting," *Journal of Development Economics*, Elsevier, vol. 85(1-2), pages 81-93. First circulated 2004 (NBER WP 10826).

Lawrence, Robert (1996): *Regionalism, Multilateralism, and Deeper Integration* (Brookings Institution, Washington DC).

Levy, Philip, 1997, "A political-economic analysis of free-trade agreements," *American Economic Review*, Vol. 87, pp. 506–519.

Low, Patrick (1993). *Trading free: The GATT and US trade policy*, Twentieth Century Fund Press, New York.

Ludema, Rodney, Anna Maria Mayda, and Prachi Mishra (2010). "Protection for Free? The Political Economy of U.S. Tariff Suspensions", June, EIIT Working Paper.

Markusen, James (2000). "Foreign Direct Investment", CIES Working Paper No. 19.

McLaren, John (2002). A theory of insidious regionalism. *The Quarterly Journal of Economics*. 117:571-608

METI (2005). White Paper on International Economy and Trade 2005, Tokyo.

Moser, Peter (1990). The political economy of the GATT, University of St.Gallen PhD thesis. St. Gallen.

O'Rourke, Kevin and Jeffery Williamson (1999). "Transportation revolutions and commodity market integration", Chapter 3 in Globalisation and History, MIT Press, Cambridge.

Ornelas Emanuel (2005c).. Endogenous free trade agreements and the multilateral trading system. *Journal of International Economics* 67:471-97

Ornelas, Emanuel (2005a). "Trade creating free trade areas and the undermining of multilateralism," *European Economic Review*, vol. 49(7), pages 1717-1735, October.

Ornelas, Emanuel (2005b). "Rent Destruction and the Political Viability of Free Trade Agreements," *The Quarterly Journal of Economics*, vol. 120(4), pages 1475-1506, November.

Panagariya, Arvind (2000). "Preferential Trade Liberalization: The Traditional Theory and New Developments," *Journal of Economic Literature*, vol. 38(2), pages 287-331.

Richardson, Martin (2001) Unilateral Liberalisation in a Multilateral World, University of Otago Economics Discussion Papers No. 0102, February 2001.

Roesseler, Freider (1978). "The rationale for reciprocity in trade agreements under floating currencies", *Kyklos*, Volume 31, Issue 2, pages 258–274, May.

Rostow, Walter (1960). *The stages of Economic Growth: A non-communist manifesto*, Cambridge University Press.

Saggi K, Yildiz HM. 2009. Bilateralism, pure multilateralism, and the quest for global free trade. Work. Pap., Southern Methodist University

Sally, Razeen (2008). "Globalisation and the Political Economy of Trade Liberalisation in the BRIICS" pdf draft, http://www.ecipe.org/people/razeen-sally.

Sauvant, Karl and Lisa Sachs (2009). *The Effect of Treaties on Foreign Direct Investment Bilateral Investment Treaties, Double Taxation Treaties, and Investment Flows,* Oxford University Press.

Schott, Jeffrey and Murray G. Smith (1988). The Canada-United States Free Trade Agreement: The Global Impact. Washington and Ottawa: Institute for International Economics.

Summers, Larry (1991). "Regionalism and the World Trading System." *Economic Review,* Federal Reserve Bank of Kansas City, issue Nov, pages 295–302.

UNCTAD (2010). *World Investment Report, 2010*. Geneva.

Vézina, Pierre-Louis (2010). "Race-To-The-Bottom Tariff Cutting," IHEID Working Papers 12-2010, The Graduate Institute of International Studies, July.

Wonnacott, Paul (1987)."The Automotive Sector in the US-Canadian Free Trade Agreement," in Murray G. Smith and Frank Stone, eds., Assessing the Canada-U.S. Free Trade Agreement. Ottawa, Institute for Research on Public Policy, December 1987, pp. 73-77.

Wonnacott, Paul (1987)."The United States and Canada: The Quest for Free Trade", Institute for International Economics, *Policy Analyses in International Economics*, #16,

World Bank (2011). *Doing Business 2011: Making a difference for entrepreneurs*, World Bank.

Yi, S. (1996), 'Endogenous formation of customs unions under imperfect competition: Open regionalism is good,' *Journal of International Economics*, 41, 151-175.

Young, Soogil (1996). "Political Economy of Trade Liberalization in East Asia", in Jeffrey I. Schott (ed.), *The World Trading System: Challenges Ahead*. Washington D.C.: Institute for International Economics.

Richard Edward Baldwin is Professor of International Economics at the Graduate Institute, Geneva since 1991, Policy Director of CEPR since 2006, and Editor-in-Chief of Vox since he founded it in June 2007. He was Co-managing Editor of the journal Economic Policy from 2000 to 2005, and Programme Director of CEPR's International Trade programme from 1991 to 2001. Before that he was a Senior Staff Economist for the President's Council of Economic Advisors in the Bush Administration (1990-1991), on leave from Columbia University Business School where he was Associate Professor. He did his PhD in economics at MIT with Paul Krugman. He was visiting professor at MIT in 2002/03 and has taught at universities in Italy, Germany and Norway. He has also worked as consultant for the numerous governments, the European Commission, OECD, World Bank, EFTA, and USAID. The author of numerous books and articles, his research interests include international trade, globalisation, regionalism, and European integration. He is a CEPR Research Fellow.