



European
University
Institute

ROBERT
SCHUMAN
CENTRE FOR
ADVANCED
STUDIES

WORKING PAPERS

RSCAS 2016/13
Robert Schuman Centre for Advanced Studies
Global Governance Programme-212

Thinking about the performance of the World Trade
Organization: A discussion across disciplines

Manfred Elsig, Bernard M. Hoekman and Joost Pauwelyn

European University Institute
Robert Schuman Centre for Advanced Studies
Global Governance Programme

**Thinking about the performance of the World Trade
Organization: A discussion across disciplines**

Manfred Elsig, Bernard M. Hoekman and Joost Pauwelyn

EUI Working Paper **RSCAS** 2016/13

This text may be downloaded only for personal research purposes. Additional reproduction for other purposes, whether in hard copies or electronically, requires the consent of the author(s), editor(s). If cited or quoted, reference should be made to the full name of the author(s), editor(s), the title, the working paper, or other series, the year and the publisher.

ISSN 1028-3625

© Manfred Elsig, Bernard M. Hoekman and Joost Pauwelyn, 2016

Printed in Italy, February 2016

European University Institute

Badia Fiesolana

I – 50014 San Domenico di Fiesole (FI)

Italy

www.eui.eu/RSCAS/Publications/

www.eui.eu

cadmus.eui.eu

Robert Schuman Centre for Advanced Studies

The Robert Schuman Centre for Advanced Studies (RSCAS), created in 1992 and directed by Professor Brigid Laffan, aims to develop inter-disciplinary and comparative research on the major issues facing the process of European integration, European societies and Europe's place in 21st century global politics.

The Centre is home to a large post-doctoral programme and hosts major research programmes, projects and data sets, in addition to a range of working groups and *ad hoc* initiatives. The research agenda is organised around a set of core themes and is continuously evolving, reflecting the changing agenda of European integration, the expanding membership of the European Union, developments in Europe's neighbourhood and the wider world.

Details of the research of the Centre can be found on:

<http://www.eui.eu/RSCAS/Research/>

Research publications take the form of Working Papers, Policy Papers, and e-books. Most of these are also available on the RSCAS website:

<http://www.eui.eu/RSCAS/Publications/>

The EUI and the RSCAS are not responsible for the opinions expressed by the author(s).

The Global Governance Programme at the EUI

The Global Governance Programme is one of the flagship programmes of the Robert Schuman Centre for Advanced Studies at the European University Institute (EUI). It aims to: build a community of outstanding professors and scholars, produce high quality research and, engage with the world of practice through policy dialogue. At the Global Governance Programme, established and early career scholars research, write on and discuss, within and beyond academia, issues of global governance, focussing on four broad and interdisciplinary areas: European, Transnational and Global Governance; Global Economics; Europe in the World; and Cultural Pluralism.

The Programme also aims to contribute to the fostering of present and future generations of policy and decision makers through its unique executive training programme, the Academy of Global Governance, where theory and "real world" experience meet. At the Academy, executives, policy makers, diplomats, officials, private sector professionals and academics, have the opportunity to meet, share views and debate with leading academics, top-level officials, heads of international organisations and senior executives, on topical issues relating to governance.

For more information: <http://globalgovernanceprogramme.eui.eu>

Abstract

International law, political science and economics scholars are all concerned with analyzing the performance of the WTO as an organization. In this paper we focus on the objectives that these different disciplines attribute to the WTO and how performance is assessed against these objectives. The literature in all three fields is vibrant, but the focus of each discipline is often on very different dimensions of WTO performance. While this implies significant complementarity across disciplines it also suggests potential opportunity costs in foregone synergies. Even when similar phenomena are the focus of analysis, different concepts, connotations and labels makes cross-disciplinary debate less efficient or prohibits it altogether. Greater effort to promote cross-fertilization across disciplines would enrich and strengthen research on the performance of the WTO.

Keywords

International cooperation, WTO, performance evaluation, multidisciplinary analysis

JEL Classification: F02; F13; F53

A. Introduction*

The World Trade Organization (WTO) officially opened its doors in 1995 at the shores of Lake Geneva. Although it is one of the youngest major international governmental organizations, the WTO built upon customs, principles and rules that developed over nearly half a century under the loosely institutionalized platform of the General Agreement on Tariffs and Trade (GATT). Notwithstanding its short history, in the past 20 years the organization has generated significant attention from various segments of society in both developed and developing countries and has attracted 33 new members, transforming into a nearly universal organization. It has also become the focus of extensive academic research spanning a number of disciplines, most notably international economics, international law and international relations. While most experts would argue that the WTO matters and research has focused on a variety of WTO-related achievements, there are conceptual and empirical holes in WTO scholarship. In particular, significant scope exists to do more to pursue inter-disciplinary research – much of the scholarly literature is limited to specific disciplinary “silos”. One reflection of this is that the questions that are analyzed by economists, legal scholars and political scientists differ significantly. While this makes research on the WTO complementary, it also implies there are unexploited opportunities for synergies.

This paper focuses on WTO outputs and outcomes by drawing on concepts that measure the performance of the organization providing different disciplinary narratives from international relations, economics and law. These debates have often occurred in isolation of each other. A key objective is to seek further cross-fertilization across disciplines. What we find is that depending on the performance matrix used (the ‘eye of the beholder’) the assessment of how the WTO matters (performance) can vary significantly. This has implications not just for research and research programmes looking forward, but is important from an accountability and “ownership” perspective. Governments have found it increasingly challenging to explain why they negotiate trade agreements and what the impacts of trade agreements have been. This is a challenge that applies to both preferential and multilateral trade agreements (the WTO) – indeed, one consequence of the “lack of coherence” in approaches to assessing the impacts of trade agreements is that it makes it more difficult to understand the rationales for, and effects of, different types of trade agreements that are pursued by WTO Members and the critical policy question where the WTO has a comparative advantage or can add most value in the pursuit of a given objective.

B. How to think about performance

As a basic understanding, one could define performance as achieving agreed-upon objectives (Gutner and Thompson 2010). In this context three key questions arise.

First, what is to be achieved? Gutner and Thompson (2010) differentiate between macro objectives which are ambitious goals, intermediate outcomes and more narrow or process-related tasks within international organizations. As we move from specific tasks to high-impact objectives achieving the set baseline becomes more difficult. Also, in reality, set goals might be imprecise or worse ambiguous. Therefore the extent to which objectives are well defined affects the overall analysis. An important qualification here is that members of international organizations might disagree on the purpose and therefore the expected tasks to be performed. This has been described as an “eye of the beholder problem”. While trade liberalization might be a key objective for one WTO member, another member might push for constraining the ability of a large player like the United States from using trade policy

* This paper was prepared for the 2015 World Trade Forum conference hosted by the World Trade Institute and University of Bern. We are grateful to Luca Rubini, Alan Winters and workshop participants for helpful comments on the conference draft.

instruments unilaterally, or seek to protect specific industries from global competition. Therefore performance assessment might vary across states and across key constituencies.

Second, given the benchmark or baseline, the next question is how well the defined objective or task is achieved. In order to answer this question, we need to focus on the sources of performance. Gutner and Thompson (2010) differentiate between internal or external factors to the organization and social vs. material factors. Internal factors relate to the delegated means (material) or the bureaucratic culture (social) prevalent in the organization. External factors address the degree to which there are shared norms (social) or power politics (material) shaping trade policy. These factors all directly or indirectly affect overall performance.

Third, an important question is how performance would have been in the absence of certain institutional features and arrangements. Establishing counterfactuals are often a methodological challenge to studies on performance (see also Elsig and Cottier 2011). In the past, research has mainly relied on comparative case study design to analyze the institutional effects. A prominent example has been the analysis of how the WTO dispute settlement system and its forerunner institution (the GATT) have dealt with similar tricky cases (Zangl 2008). A special test for the impact of institutions is how they perform in hard times. In the case of external shocks such as the economic and financial crises, it is important to understand how existing institutions buffer against demands for increased protectionism.

In sum, measuring performance is far from trivial. It is composed of an analysis that first establishes the various baselines before mapping and tracing the key sources of performance in order to understand causal mechanisms. Finally, it often deals with methodological challenges for establishing causality speaking to the need to think carefully about counterfactuals. In the following we present three different angles to performance rooted in international relations, economics and legal scholarship.

C. An international relations angle

Most scholars in the mainstream international relations (IR) tradition consider the role of international organizations (IOs) as rather limited. However, institutions created in the trading system have been acknowledged as being more influential than institutions in many other policy fields. Already the seminal work of Keohane (1984) attributes to the GATT an important function to overcome states' reluctance to cooperate by providing information and helping induce compliance. Later, the WTO was excluded explicitly from realist-type skepticism that IOs without enforcement mechanisms are inconsequential (Downs et al. 1996). Most realist writings clearly stated that the WTO mattered as an institution as it was able to modify its members' behavior. The realist contributions were focusing more on the question who designed the rules and who controls the organization (Gruber 2000). In addition, some work has suggested that increased legalization could hurt weaker states (Drezner 2007). In particular overlapping new legal obligations create a fragmented system which allows larger states to strategically engage in forum-shopping of legal institutions. Power arguments were also tested in relation to weak states' (lack of) participation in the WTO dispute settlement system (Sattler and Bernauer 2011, Elsig and Stucki 2012).

I. So, what is to be achieved?

In order to discuss the performance of the WTO, we need to start by asking what its purpose is. While there are formal objectives listed in the treaties (e.g. in the preamble), these are often not suitable baselines for an assessment because they are either too ambitious in the sense of being beyond the control or influence of the organization or difficult to operationalize for the purpose of measurement (Elsig 2010). Moreover, if we can locate some objectives, these are subject to significant change over time and actors' support for these might weaken or strengthen. For example, the original 1947 GATT

agreement did not explicitly mention economic development (e.g. poverty reduction) as a goal to be achieved, while environmental concerns were absent as witnessed by the aim of “developing the full use of the resources of the world.” Notwithstanding differences of views on objectives and what importance they have, the multilateral trading system was created to help increase the international exchange of goods (and, following the Uruguay round, trade in services) based on the general presumption that lowering barriers to trade increases overall welfare. While the economics literature (see below) has focused very much on this dimension, the IR literature has put attention more to questions related to how the GATT/WTO institutions have served various political objectives. Below we distinguish between macro objectives and micro objectives.

A group of general objectives which are not core trade objectives, but which are important for the stability of the system, could be labelled macro goals. These include:

First, international institutions are purposefully designed to uphold the “rule of law” which is reflected in rules which serve to protect in particular the weaker states and to control the abuse of power (Grant and Keohane 2005). This is particularly true for the WTO and its dispute settlement system. One of the main incentives to negotiating a highly legalized system to support the implementation of new sets of rules was to tame US unilateralism (Elsig and Eckhardt 2015, Pelc 2010). Therefore, a baseline objective can be derived around a shared understanding held by a large group of contracting parties to constrain US unilateralism and to demand the US to use the new dispute settlement tools offered by the WTO.

Second, another potential objective is the ability of multilateral organizations to accommodate new powerful states. As some earlier GATT trade rounds were designed to integrate the EC and the emerging economies (e.g. Brazil and India) into the multilateral system, the WTO helps to manage changing trade power dynamics driven by China’s sustained high growth rates. China joined in 2001 after having gone through the accession procedures. More recently, WTO Members and Russia agreed on Russia’s terms of membership. Integrating these states into the existing system could certainly be seen as a major objective for international cooperation. The baseline here could be defined as the new entrants playing constructively the game (following the rules and rulings, providing leadership in ongoing negotiations).

Third, good international institutions are those that also work well in times of distress. Or put differently, an organization’s contribution to problem-solving could prove particularly important in times of crises. In the context of the GATT/WTO system, one can consider the economic and financial crises which led to calls by domestic actors for protecting national markets as a particular challenge. The question is whether the WTO can mitigate the negative effects that result from increasing protectionism.

Fourth, focusing on the weakest of the system, another potential baseline is to integrate less developed countries (LDCs) into the system and provide support in negotiations, dispute settlement, or implementation-related matters. For the legitimacy of the WTO (Elsig 2007), it is important to allow for differentiated treatment with a view of generating benefits for all WTO members. While LDCs were in the past mostly excluded from concessions, more recently LDCs have been called upon to engage and take on commitments. The baseline would consist of measuring the success of such integration and moving from exclusion to increasing involvement.

While the above objectives are situated at the macro level and primarily define the success of international organizations such as the WTO, one can also consider less ambitious, institutional and procedural goals and functions that support the realization of the macro objectives. Four “services” or institutional platforms stand out in the context of the WTO:

First, the simple provision of a negotiation platform where WTO members can regularly meet, exchange concessions and explore new regulatory solutions to ongoing challenges. In the end of the day, the baseline is the amount of outputs that are generated through these negotiations with notable

effects for market liberalization and addressing “disguised” protectionism. Currently, the assessment of the outcome of the negotiation platform is straightforward given the lack of real progress on new agreements with very few exceptions (e.g. the expansion of the information technology agreements, the agreement on trade facilitation).

Second, an important function relates to regime management activities. In order to support the implementation of the agreements, the organization assists through peer reviews, discussion fora and the active use of transparency mechanisms. Different tools stand out. One is the “member-driven” work in the various Committees which allows for clarification and tabling concerns, learning about best practice and can potentially lead to the elaboration of new regulatory initiatives (Lang and Scott 2009). Another is the Secretariat-driven assessment of national trade policies (e.g. Trade Policy Review Mechanism). This activity helps provide more information and could potentially turn into a naming and shaming exercise. Overall, the question is how well the system supports the implementation of the obligations (1st order compliance) and how Committees may contribute to design new initiatives.

Third, support by the organization to WTO members through its technical assistance and capacity building arm. Are we witnessing improvement in Members’ capacity to participate in the system? This supports directly the integration of LDCs into the system (macro goal). However, assessing technical assistance work is far from straightforward and to our knowledge no field experiments have been conducted to isolate the potential impact of the WTO from other factors.¹

A fourth function is the provision of ‘affordable’ and equal access to the dispute settlement mechanism. Can affected members launch legal proceedings given power asymmetry, lack of direct support by stakeholders or lack in legal capacity? These objectives support weaker actors and help them integrate into the system. Some research has suggested continued lack of legal capacity (Kim 2008, Busch et al. 2009) as well as power asymmetry which leads to weak states not launching many complaints against more influential states (Elsig and Stucki 2012, Sattler and Bernauer 2011).

II. Sources of performance

Whatever the exact macro or micro objective that is studied, an important next step is to attempt to locate the key internal and external factors as well as social and material factors that influence performance. Which one of these multiple sources is often a function of the aspired objective? More salient and politicized goals (or where disagreement among Members on the direction is significant) may be more likely affected by material interests and external factors, whereas less contentious objectives and procedural goals might be significantly shaped by the institutional set-up and social factors.²

To provide an example of how these factors come into the picture, let us focus on the stalled negotiations. How important are different factors for reaching the objective of concluding the Doha Round? A key external factor has been the emergence of a multi-polar trade world which creates a larger set of “material” interests that diverge over the Doha outcome. Increased interest competition translates into varying expectations about the content of the Round. In other words, while few would dispute the importance of wrapping up the round, the “eye of the beholder” problem as to what should be “in the package deal” leads to hard bargaining. Internal institutional challenges complicate the finding of agreeable solutions. That the decision-making mode is largely based on consensus-decision

¹ It would be interesting to study the effects of participation in training activities by comparing groups that were exposed to this „treatment“ and groups that have not participated in training. A major challenge here is to control for other actors’ influence on technical assistance and training, as the WTO is just one of many providers of such services.

² However, we can think of alternative interaction effects, for instance internal and material interests (e.g. the interest of bureaucratic power maximization by IO staffers) or external and social factors (e.g. Washington Consensus affecting the work of the WTO).

making makes the increased heterogeneity of interests more visible and creates an internal collective action problem. This “institutional milieu” of the WTO (Elsig 2010) is further complicated by a socially constructed internal understanding that this organization does not vote, although the treaty would allow to take recourse to voting. Thus, in this example external and material interests are key, but the lack of performance can only be fully understood when focusing on internal norms and rules as well.

Collective action problems might in turn loom less large in WTO activities where norms are shared and material interests less conflicting. This might explain why WTO Members have agreed on a deal at the WTO Ministerial Conference in Bali on Trade Facilitation. It however might also be a result of smart design in which both developed and developing countries expect to reap benefits.

III. Assessment

Based on the above conceptualization of performance, the question whether micro goals have been achieved cannot be easily answered without more in-depth causal analysis. We certainly witness areas of non-performance in relation to the negotiation function of the organization. Here the eye of the beholder problem related to the content of treaty outcomes looms large. By contrast, we also observe progress in terms of dispute settlement. Few would have expected the WTO dispute settlement system to play such a prominent role when it was created (see also Section E below). As to other micro goals, such as regime management and technical assistance and capacity building, the assessment is rather positive with room for improvement (e.g. using the true potential of committee work).

If we focus on macro goals, most of these look *prima facie* as achieved to a large degree. The US has taken less recourse to unilateral actions and has been an active user of the dispute settlement system. Overall there is little evidence that the system has not moved toward more law (and therefore to constrain the abuse of power). The integration of China in terms of dispute settlement seems to have worked well in light of avoiding “trade wars”; the leadership role in negotiations might be another story. For the integration of Russia, it is too early to make a judgment call. Finally, the system was not damaged through the financial and economic crises. Whether this has been a result of careful retention by WTO members by not bringing many more new cases or an actual protectionism-taming-effect of the WTO’s legal system remains unclear (see Baccini and Kim 2012). A counterfactual method could be employed in studying this particular event.

D. An economics perspective

Economists approach the question of assessing the objectives and performance of the WTO through a variety of lenses. What follows focuses on two dominant viewpoints/approaches. The first centers on understanding the rationale for trade agreements – the incentives governments have to cooperate. The second lens is primarily empirical, the aim being to assess the effects of WTO membership on individual countries and the impact of the organization as a whole. Research in this vein centers on the design and results of negotiations aimed at reducing trade distorting policies. The focus is on performance in the sense of understanding the economic implications of both what is being (might be) negotiated – *ex ante* analysis – and what has been agreed – *ex post* analysis.

Empirical economists have assessed the impact (‘performance’) of the WTO using a range of metrics, including the volume (value) of total trade flows, the structure and composition of trade flows (e.g., changes in the variety of goods traded; increases in the number of trading partners, etc.) and economic welfare (real income). Such empirical research is both “macro” in the sense that it does not consider the specifics of any of the rules but simply analyzes whether a “WTO effect” can be discerned in the data, and “micro” in the sense that the focus of analysis is on a specific area or set of

disciplines – antidumping, agriculture, intellectual property, services, safeguards, technical barriers to trade, trade facilitation, the DSU, etc.³ Whether ‘macro’ or ‘micro’, it is usually taken as given that WTO disciplines and commitments are (have been) implemented and analysis tends to abstract from implementation costs in the narrow sense (see e.g., Finger, 2007).

I. Objectives of WTO Members

There are two strands of theorizing in the economics literature about the ‘macro’ goals of countries that engage in trade agreements. One revolves around the terms of trade: the premise is that countries negotiate so as to reduce the negative terms-of-trade effects generated by trade policies of partner countries (Johnson, 1953-54; Bagwell and Staiger, 2002).⁴ A consequence of this view of trade cooperation is that agreements will (need to) be among countries that can affect their terms of trade. Small countries will generally have much less power to affect their terms of trade, although if products are differentiated even a small state will have some power to influence them.⁵

Another strand of the economic literature (e.g., Finger, 1979; Tumlrir, 1985; Maggi and Rodriguez-Clare, 1998; Ethier, 2007) argues that trade agreements permit governments to adopt welfare-enhancing policies that are not (politically) feasible otherwise. The premise here is that governments are conservative in the sense that they put greater weight on prospective losses for groups in society than on the expected gains from liberalization. More generally, governments have incentives to impose or maintain protection because this raises the incomes of the groups from which they derive political support. By committing to certain rules that constrain policy space a government can make its reforms more credible: officials can tell interest groups seeking the (re-)imposition of trade policies that doing so will violate the nation’s WTO commitments and generate retaliation by trading partners. Trade agreements ‘work’ because they change the domestic political economy of trade policy by mobilizing export interests to oppose lobbying for protection by import-competing interests because they would then lose access to foreign markets. By affecting the balance of political support a more liberal trade policy stance becomes optimal.

The terms-of-trade and commitment-cum-political economy approaches generate clear goal posts from a performance perspective: both imply that the outcome of cooperation is a mix of liberalization (lower average trade barriers) and rules that constrain the ability to re-impose barriers. The different theories also all imply that an effect of the WTO will be to increase trade relative to a counterfactual where no agreement had occurred – because of the reduction in barriers and other trade-distorting policies, and the reduction in policy uncertainty.

Views differ on how strong the evidence is for the presumed underlying rationale for trade cooperation and this is an active subject of research. However, given that both theories imply that an outcome of cooperation is trade liberalization and policy bindings that reduce uncertainty for traders, common measures of WTO ‘performance’ include (i) progress in further reducing trade barriers and agreeing to new policy disciplines; (ii) the extent to which domestic political economy pressures to re-impose protection were contained in times of stress and in response to shocks; and more generally (iii)

³ There is a huge literature on the many different WTO agreements and the effects, both normative and positive, of the policy disciplines that are embodied in each agreement. See, for example, the 21 volume series *Critical Perspectives on the Global Trading System and the WTO* (Edward Elgar Publishing) for a collection of what leading scholars regard as core contributions on each of the main WTO agreements and activities.

⁴ As such negative externalities can be created by domestic policies as well, the terms of trade theory also offers guidance regarding what types of policies could be the subject of multilateral negotiations and cooperation.

⁵ Bagwell and Staiger (2011). Of course, governments of small countries that cannot affect the terms of trade still have an interest in being a member of the WTO because exporters will benefit from the tariffs that larger countries negotiate reciprocally with one another and then extend to all members under the MFN rule.

trade effects—whether the disciplines and operation of the WTO are associated with changes in trade and specialization.

II. Goals and performance of the WTO: Three metrics

1. Reducing trade barriers and agreeing to new policy disciplines

The basic stylized fact here is well-known: to date the WTO's performance in delivering new policy commitments has been very weak. This does not require economic analysis, of course, in that no new market access agreements were negotiated between 1995 and 2015. The Doha Development Agenda has been moribund since 2008 and the large players that historically have been the mainstays of multilateral trade cooperation (US and EU) have shifted their main focus to the conclusion of so-called mega-regional agreements (TPP, TTIP). Indeed, there has been a general shift towards preferential trade agreements (PTAs): some 400 PTAs have been notified to the WTO in the last 20 years. Much of the liberalization that has occurred since 1995 has been associated with either unilateral reforms or the implementation of PTAs.

An interesting question is to what extent this is consistent with the 'macro' goal of reducing barriers to trade. Given that average levels of protection have continued to fall throughout the post-1995 period, and that the WTO explicitly allows for members to negotiate PTAs, it is not as clear as is often argued that the WTO has failed to deliver. It has not proved possible to conclude the Doha Round and agree to new rules in areas such as agricultural subsidies, but at the same time governments have signed numerous preferential trade agreements and reduced average levels of protection. Moreover, it should be recognized that in key areas that economic research suggests will deliver large gains for developing countries – most notably trade facilitation (where estimated potential gains exceed what was proposed in Doha for reducing tariffs—e.g., Hoekman and Nicita, 2011) – WTO members did conclude an agreement.

2. Managing political economy pressures and preventing backsliding

Does the WTO improve national trade policy processes and outcomes? Here the answer is largely positive, although there is no consensus. At the national level there is evidence of a positive effect of the WTO on the 'quality' of trade policy. In part this is reflected in the fact that although numerous, the number of disputes are much less than what would be predicted based on the number of bilateral interactions between countries (Horn et al, 1999; Sattler and Bernauer, 2011). In part this has been demonstrated by case studies of specific countries. Examples include the work of Mike Finger and coauthors (Bacarat et al. 2013; Finger and Nogues, 2006) showing how WTO disciplines can be and have been used to manage and deflect pressures for protection by powerful interest groups.

One measure of performance on this front is what happened (did not happen) following the 2008 financial crisis in the US and the EU. In contrast to the late 1970s and early 1980s, the closest recent analogue where a major shock (OPEC price rise) sent the global economy into a tailspin, when governments imposed a plethora of trade restrictions, including quotas and 'voluntary' export restrictions, there was not a significant rise in average levels of protection post-2008.⁶ For example, Gawande, Hoekman and Cui (2015) find that WTO disciplines were a factor, although they also point to changes in the incentives to use trade policy: once countries are highly integrated into international production networks they will not benefit from trade protection. But insofar as the rise of global value chains was a factor in restraining protectionism, to a significant extent this in itself can also be

⁶ While views on this differ among economists – e.g., Evenett and Fritz (2015) argue there has been a significant increase in the use of trade-distorting policies – the majority view among economists is that average levels of protection did not increase significantly. See e.g., the contributions in Hoekman (2015).

attributed to the WTO: the open, rules-based multilateral trading system underpinned the shift by companies towards greater specialization, cross-border investment and international production networks. Another measure of the performance of the WTO in this area is the absence of a large scale increase in disputes alleging violations of WTO commitments post-2008 crisis (see also Section E below). As discussed subsequently, one of the ‘micro’ functions that played a supporting role in this regard are the WTO’s various transparency mechanisms.

3. Trade effects of the WTO

As noted previously, one measure of the attainment of both of the broad macro goals of the WTO that are distinguished in the economics literature is whether the WTO has an effect on countries’ trade patterns and performance. The most straightforward measure is whether the WTO overall or membership of a country in the WTO increases trade volumes and the structure of trade (lower barriers and less uncertainty should increase investment in tradables and allow exploitation of comparative advantages that will be reflected in changes in specialization).

The best evidence from a statistical perspective – in terms of identifying a ‘WTO effect’ – comes from studies of the effects of WTO accession. Countries that acceded to the WTO after 1995 differ from ‘original’ WTO members in that they were forced to go through a rigorous and demanding process of review and scrutiny, and were required to do much more to bolster and reform trade policy institutions than GATT incumbents. Rose (2005); Subramanian and Wei (2007) and Tang and Wei (2009) find a positive trade impact for countries that joined the WTO after 1995. Tang and Wei (2009) attribute this positive effect to improved economic governance and argue that WTO accession acts as a partial cure for weaknesses in the investment climate in many acceding countries. The mechanisms that are at play here span various elements of the WTO, including greater policy certainty associated with bindings and transparency. Reducing uncertainty is an important dimension of the WTO and trade agreements more generally (see, Francois, 1997; Francois and Martin 2004; Handley 2014 and Limão and Tovar 2011).

There has been debate in the economics literature whether positive trade effects apply more generally. Rose (2004), most notably, undertook a gravity model analysis of the determinants of bilateral trade flows of GATT/WTO members for a 50 year period and failed to find a distinct impact associated with GATT/WTO membership. However, Subramanian and Wei (2007) find that the WTO does increase trade – by some 40 percent on average – but only for those countries that participated in the process of reciprocal exchange of trade policy commitments. They do not find such a positive trade effect for most developing country members of the GATT, which they attribute to the pursuit of special and differential treatment which implied that most of these countries did not make changes to their trade regimes. Similarly accounting for this selection bias, Liu (2009) and Dutt et al. (2013) find a strong role for the WTO in increasing the extensive margin of trade; the latter conclude that WTO membership increases the extensive margin by 30 to 40 percent.

Another example of a WTO agreement that should allow for clear identification of impacts is the Agreement on Government Procurement (GPA). This applies to only a subset of countries, thus permitting comparisons between signatories and non-signatories. Research finds no effects on the import behavior of participating governments. Shingal (2015) analyzes procurement sourcing over time in Japan and Switzerland, and finds that GPA membership has no independent effect. Shingal (2011) concludes that the share of services contracts awarded to foreigners declined over time for these two countries, as opposed to rising. Rickard and Kono (2013) assess the effects of the GPA and 43 PTAs that include procurement, and conclude the agreements have no impact on foreign sourcing. However, here again there may be country-specificity at work. Fronk (2014), focusing only on agreements negotiated by the US finds a statistically significant positive effect of bilateral procurement agreements on procurement behavior. Thus, while the WTO overall has had the expected positive trade impact, this is not the case across the board.

III. Sources of Performance

The determinants of the performance of the WTO in achieving the broad ‘macro’ goals that are the focus of economics research are varied. External events – not just events such as the 2008 crisis/recession or the 1998 East Asian financial meltdown but also major structural changes such as the rapid growth sustained by China that led to a major rebalancing of the world economy – clearly are a factor, as are decisions by WTO members whether and how to employ outside options to pursue the macro goals. The intensity with which WTO members have used PTAs increased in the WTO’s first 20 years and it is likely that this trend will continue. It is important to consider that the WTO makes explicit provisions for the negotiation of PTAs between its members. The WTO membership has a choice where and how to pursue trade cooperation. The revealed preference has been to pursue parallel tracks: the Doha negotiations and plurilateral initiatives as well as PTAs. Has a proliferation of PTAs led to significant trade diversion and negative spillovers for non-members? While the jury is still out on this question, empirical economic research suggests that PTAs have not had anywhere near the negative effects they might have had and that on balance they have been associated with a general trend of greater openness to trade globally.⁷

PTAs do not address many of the core issues that are of concern to the majority of the WTO membership and that are on the table in the DDA – most notably trade-distorting agricultural policies and subsidies. They are not a substitute for a comprehensive multilateral agreement that goes beyond what was negotiated in the 1990s (the Uruguay Round). An internal factor that many argue plays a role in the limited achievements in new rule-making in the WTO is how the negotiating agenda is determined and limited scope to adjust this over time. Numerous economic analyses suggest that the ‘landing zone’ associated with the DDA is both small and perhaps more important, that the distribution of associated economic benefits and losses is quite asymmetric (Decreux and Fontagné, 2015; Laborde and Martin, 2015; Martin and Messerlin, 2007). These are issues where there is substantial scope to apply both quantitative modeling and analysis used in other parts of economics (e.g., the theory of clubs and fiscal federalism), as well as a clear need for multi-disciplinary research strategies to identify areas of complementarity between PTAs, small group cooperation in the WTO, and universal membership agreements.

A function of the WTO that has played an important role in sustaining cooperation – preventing backsliding; avoiding disputes – are the various transparency mechanisms. This played a positive role in the post-2008 period in raising external visibility of national trade policy measures and signaling that actions of governments would be scrutinized. Of course, determining to what extent this led to fewer action being taken is difficult to determine – there is no clear counterfactual. This is one subject where there is a need for multi-disciplinary analysis, e.g., detailed case studies of deliberation within governments and in parliaments. The same is true for assessments of the role that WTO commitments and disciplines play in the way that governments respond to pressures from domestic constituencies for economic assistance.

WTO rules, processes and interactions are just one mechanism through which government policy is affected. The economic research on the trade and growth effects of the WTO demonstrates that these effects may be indirect as well as direct. The indirect channel may well be more important, but difficult to identify. Thus, engagement in the WTO may help to stimulate a focus on improving the business environment and investment climate more generally. The need to develop a capacity to implement agreements and to participate in dispute settlement may have spillover effects in other areas – e.g., the negotiation of investment agreements, or a reconsideration of approaches towards non-trade economic regulation. The WTO cooperates and interacts more closely with a large number of

⁷ In a survey of the empirical literature Freund and Ornelas (2010) conclude that overall PTAs have not had the adverse effects that many had expected them to have, such as significant trade diverting effects, because they were associated with a period in which countries liberalized more generally. Estevadeordal, Freund and Ornelas (2008) conclude that recent vintage PTAs have been accompanied with a reduction in MFN barriers to trade.

international agencies and organizations than the GATT did. An example is the work on ‘aid for trade’ which has brought the WTO into the ambit of the development community and raised awareness of trade capacity needs in the development world. This may well have had significant payoffs in attaining ‘macro’ goals. Documenting what should be attributed to the operation of the WTO is a difficult challenge.

IV. Assessment

Space constraints have limited this section to a very selective and partial discussion of the economic literature that is relevant to an assessment of the performance of the WTO. However, on balance this literature suggests that if the focus is on the three objectives identified above the organization has performed relatively well. The evidence indicates that the WTO has generated lower barriers to trade, sustained these in times of stress, and generated more trade. Insofar as greater trade is associated with higher incomes and a reduction in poverty, the WTO has increased world welfare – which is of course the primary high level goal identified in the WTO Preamble. Of course, there are serious attribution problems associated with any empirical estimates of how much the WTO has contributed to global welfare (see, e.g., Anderson, 2014), and matters are complicated by the fact that the WTO builds on an acquis that was negotiated over a long period of time, including the GATT years. Looking forward it is clear that future assessments of performance will depend on whether and how WTO members are able to multilateralize what they achieve through PTAs and their willingness to do so.

E. International law perspectives

In legal scholarship, performance studies have focused on the effectiveness of international courts and tribunals or the dispute settlement mechanism set up in the context of certain IOs. With the considerable increase in the creation of international courts and tribunals, especially since the 1990s, the question of how these international adjudicators have performed has attracted significant attention. The puzzle of how international adjudicators can be effective without genuine police or enforcement powers remains central. In the WTO, this led to a focus on the performance of WTO dispute settlement.

The predominant approach to assessing the performance of WTO dispute settlement is less than satisfactory. Most legal scholars and WTO insiders be it from the WTO secretariat or diplomatic missions of WTO member countries have focused on either (i) usage rates, that is, how many disputes have been filed at the WTO, or (ii) compliance with adverse rulings issued by WTO panels or the Appellate Body (Wilson 2007; Davey 2009; Hughes 2012).

That usage rates or compliance with court judgements are at best limited proxies for performance should, however, be obvious. Indeed, at annual updates on WTO dispute settlement, chairs of the WTO Dispute Settlement Body (DSB) have invariably referred to both low and high usage rates as evidence of success.⁸ When usage rates are *low*, the message is that WTO countries comply with WTO rules and jurisprudence, so there are no disputes and the system works. When usage rates are *high*, insiders congratulate themselves by stating that a high number of disputes demonstrates trust in, and effectiveness of, WTO dispute settlement. Either of these views can be correct (low usage rates, for example, can, indeed, be evidence of high levels of compliance, or of complete ineffectiveness or unattractiveness of the system) but they cannot be both correct at the same time.

⁸ See Reflections from Outgoing Chairs on DSB Developments, 2010-2014, available at https://www.wto.org/english/tratop_e/dispu_e/dispu_e.htm#dsb. See also WTO Director-General Azevedo’s statement, in reaction to WTO disputes reaching 500 mark: “This shows that the WTO’s dispute settlement system enjoys tremendous confidence among the membership, who value it as a fair, effective and efficient mechanism to solve trade problems” (WTO 2015).

Similarly, there is no doubt that the rate of compliance with adverse WTO rulings by panels and the Appellate Body is impressively high (more than 80 %, Davey 2009:119; according to the WTO website, “around 90%”, WTO 2015). Still, this alone tells us little, if anything, about overall compliance with WTO rules (the WTO disputes actually filed and ruled upon may only be the relatively uncontroversial tip of the iceberg of total number of violations out there), nor even about the actual compliance pull of WTO rulings (WTO rulings may be complied with because they impose relatively low burdens, because losing states see it in their own interest to do so anyhow or for reasons external to the WTO such as unilateral threats to pull-back development aid). As Shany puts it, “a *low-aiming* court, issuing minimalist remedies, may generate a high level of compliance but have little impact on the state of the world” (Shany 2012:227). The European Court of Human Rights, for example, boasted until recently about an almost perfect compliance rate. However, once the remedies provided became more intrusive -- not just declaratory statements that a violation had occurred or relatively conservative monetary compensation, but also orders to reopen faulty legal proceedings or to adopt broad legal or policy reforms -- compliance rates significantly declined (Shany 2012:263-4).

Similarly, high compliance with WTO judgements may be related to the relatively limited remedies provided: only prospective compliance in the individual case; no retroactive damages for past harm caused or obligation to reform *v-à-v* all WTO members. Conversely, a certain degree of non-compliance with WTO judgements does not necessarily reflect badly on WTO performance. As Raustiala points out, international rules or rulings “can be effective even if compliance with them is low. If a legal standard is quite demanding, even widespread failure to meet it may still correlate with observable, desired changes in behavior” (Raustiala 2000:394). Also, less-than-perfect compliance may be good rather than bad for the WTO system: It provides an escape valve or limited “exit option” for countries to accommodate exceptionally strong political or economic interests. This “exit option” in the face of exceptional circumstances may, indeed, explain why WTO members were able to sign onto the more legalized WTO dispute settlement system in the first place. Without it, they may also be less willing to consent to new agreements (Pauwelyn 2005). On this view, attempting to ratchet up compliance levels with WTO rulings to 100% may backfire and harm the system (e.g. certain countries may leave the WTO) rather than enhance its performance.

I. Performance: A new approach

Non-WTO specific studies on the performance of international courts and tribunals take a more structured approach, along the lines we suggest in section B above (establishing the benchmark, isolating key sources of performance and thinking hard about potential counterfactuals).

Rather than focusing on “outputs” such as number of disputes filed or rulings issued, Shany (2012) proposes a “goal-based approach” -- assessing “outcomes” rather than “outputs” -- to assess the performance of international courts: “in order to measure the effectiveness of an international court using this approach, one has to identify the court’s aims or goals—that is, the desired outcomes that it ought to generate—and ascertain a reasonable time frame for meeting some or all of these goals” (Shany 2012:225). At a minimum, this approach involves “goal identification, outcome assessment and establishing causation”. A more comprehensive analysis would “also refer to structural and procedural indicators in order the gauge outcomes better and to diagnose root causes for underperformance” as well as “the cost-effectiveness and efficiency of specific goal-attainment strategies” (Shany 2012:254-5).

Shany recognizes that an international tribunal may have different goals and distinguishes on the basis of their source (goals set by external constituencies *v.* actors within the system), hierarchical level (ultimate ends *v.* intermediate goals) and method of articulation (goals explicitly identified *v.* implicit or unstated goals). He also identifies four generic goals that most international courts are encouraged to achieve: (i) “promoting compliance with the governing international norms (primary norm compliance)”, that is, in the WTO context, the enforcement of WTO commitments or seeking

compliance by WTO members with WTO norms, not only through disputes filed but also by means of changing or influencing state behavior “in the shadow of the law”, (ii) “resolving international disputes and specific problems (dispute resolution or problem solving)”, in the WTO sphere, avoiding escalating trade wars or settling trade disputes between two specific WTO members, (iii) “contributing to the operation of related institutional and normative regimes (regime support)”, in the WTO, for example, to contribute to the WTO’s overall goal of trade liberalization, and (iv) “legitimizing associated international norms and institutions (regime legitimization)”, in the WTO, to confer legitimacy on, for example, the states and government officials that established the WTO or to, more broadly, advance the rule of law in international relations.

Along similar lines, Helfer (2014) identifies four different dimensions of performance of international tribunals: (i) “case-specific effectiveness”, a measure closely linked to the level of compliance with court rulings, (ii) “erga omnes effectiveness” or the effect of court rulings on the behavior of all states subject to the court’s jurisdiction (not just the disputing parties), (iii) “embeddedness effectiveness” or “the extent to which [international courts] anchor their judgments in domestic legal orders, enabling national actors to remedy potential treaty violations at home and avoid the need for international litigation” and (iv) “effectiveness in developing international law” or “norm-development effectiveness”, that is, how court rulings “contribute to building a body of international jurisprudence” (Helfer 2014:466).

Applying these insights to an assessment of the performance of WTO dispute settlement, five steps can be envisaged: (i) goal identification, (ii) measurement, (iii) establishing causation, (iv) assessing cost-effectiveness and (v) considering explanations or indicators explaining good or bad performance (leading to both diagnosis and cure).

1. Identification

On goal identification, it is clear that the DSU may, for different people, have different goals. For trade diplomats, the main goal may be the avoidance of escalating trade wars and the settlement of trade disputes based on rules rather than economic power (Shany’s “dispute resolution or problem solving” function; Helfer’s “case-specific effectiveness”). For private traders, the core goal may rather be to enhance market access in foreign countries by efficiently removing WTO inconsistent trade barriers (Shany’s “primary norm compliance” and “regime support” functions). For academics and international law practitioners, in turn, the main benchmark may be the quality of legal or economic reasoning of WTO rulings or the system’s contribution to world trade law or, more broadly, public international law (Helfer’s “norm-development effectiveness”). Depending on the goal identified, the WTO dispute settlement system may then be highly successful (e.g. in avoiding escalating trade wars, achieving compliance with actual court rulings or developing a stable jurisprudence), or deeply disappointing (e.g. in swiftly opening foreign markets for private traders, significantly reducing the total pool of trade protectionist measures or basing WTO rulings on solid economic evidence or reasoning).

As noted earlier, high levels of compliance with court rulings (“case-specific effectiveness”) does not necessarily equate with high levels of overall norm compliance (“primary norm compliance”), see also Elsig (2016). In the WTO, for example, it is interesting to note that in the first 20 years (1995-2014), 488 requests for consultations were filed, leading to 154 panel reports and 96 Appellate Body reports; with the majority of disputes settled or not proceeding, only 18 disputes where the remedy of retaliation was authorized and close to 90 % compliance with court rulings. Yet, the distribution of cases across the different WTO agreements is very uneven: out of 488 requests, 387 invoked GATT, 107 the Anti-Dumping Agreement and 104 the Subsidies Agreement; only 23 the GATS and 34 the TRIPS Agreement.⁹

⁹ DSU Statistics made available by the WTO Legal Affairs Division.

Low usage rates for GATS and TRIPS may indicate high levels of compliance with those agreements; it may also point at ineffectiveness of the system in those sectors or reluctance by WTO members to enforce commitments in these areas. In any event, the end result is that WTO jurisprudence is highly developed in some areas (for example, national treatment and trade remedies) but underdeveloped or non-existent in other fields (for example, large parts of GATS and TRIPS, or GATT Article XXIV on regional trade agreements). The absence of disputes or jurisprudence under GATT Article XXIV is certainly not explained by the absence of free trade agreements. It may be explained by compliance by members with the conditions under Article XXIV or, more likely, the adage that “people who live in glass houses do not throw stones”, that is, many WTO members may violate GATT Article XXIV; hence, they are unlikely to sue another country on the issue for fear of a ruling backfiring against themselves.

WTO dispute settlement may thus perform better in some functions than others (e.g. more “case-specific effectiveness” than “primary norm compliance”). For one and the same function (“norm-development effectiveness”), it may perform better in some fields (national treatment) than others (GATT Article XXIV).

When it comes to Helfer’s “embeddedness effectiveness”, for example, there can be little doubt that WTO dispute settlement has (predictably) performed rather badly: for lack of direct effect of WTO law in the national legal systems of the major trading nations and the absence of any requirement to first exhaust domestic remedies before resorting to WTO dispute settlement, domestic courts have played almost no role in interpreting and enforcing the WTO treaty. Conversely, to the surprise of some, WTO dispute settlement and especially the WTO Appellate Body has played a major role in the elaboration and development of not only WTO law but also general principles of public international law (e.g. rules on treaty interpretation, burden of proof or good faith), thereby performing well in terms of “norm-development effectiveness”.

Crucial in this respect is to keep in mind also that bad or good performance under a particular benchmark is not by definition bad or good normatively speaking. For many WTO insiders or diplomats, for example, neither “embeddedness” nor “norm-development” effectiveness is particularly desirable. Rather than delegating power to domestic courts or WTO panels and the Appellate Body, they may want to keep enforcement and lawmaking powers in their own hands. What is admirable norm-development for one person is unacceptable judicial activism for another. For some the DSU is there to boost and facilitate the enforcement of WTO commitments, for others, in contrast, it is there to limit and control enforcement especially unilateral enforcement by the most powerful countries e.g. by putting a cap on authorized retaliation.

2. Measurement and causation

Once goals and benchmarks identified and unpacked (step 1), as complex as this may be, the second and third steps of measurement and causation are arguably even more difficult. Measuring “primary norm compliance” is particularly daunting: How many violations of WTO law are out there and how many would there have been “but for” WTO dispute settlement? Past research has worked with proxies to get an approximate measure of primary norm compliance, usually not going beyond single country studies (e.g. Davis 2012 on the US). These studies rely on government reports which list questionable import and behind-the-border measures imposed by trading partners (see Elsig 2016).¹⁰ Gauging “case-specific effectiveness” may be easier: Of all trade disputes raised, how many are effectively resolved thanks to WTO dispute settlement? However, even there causation may be problematic: How does one demonstrate that a country withdrew its trade restriction, settled a case or complied with a WTO ruling because of WTO dispute settlement and not because of factors external to the WTO such as a change in government in the defendant country or military or development aid

¹⁰ See also de Bièvre (2015).

threats uttered by the complainant? Comparisons between WTO dispute settlement and its predecessor under the GATT or other international tribunals can be useful but have limits (see Zangl 2008). Pointing out that over the same period of time more cases have been filed to the WTO than the GATT (500 requests for consultations in the WTO in little more than 20 years compared to 300 disputes over a period of 47 years of GATT, WTO 2015) does not necessarily mean that the WTO performed better. As discussed earlier, more cases may be due to more norm violations out there (hence a lower level of “primary norm compliance”). The WTO may also have more cases than the GATT or the International Court of Justice because more countries fall under its compulsory jurisdiction or the WTO covers a broader set of issues.

3. Cost-effectiveness

Cost-effectiveness is particularly important but often overlooked. It may well be that WTO dispute settlement is effective in settling trade disputes or maybe even in substantially changing state behavior away from trade protectionism. However, is WTO dispute settlement the most cost-effective means to achieve whatever results found as compared, for example, to peer review monitoring or softer, more collaborative mechanisms of ensuring implementation? The “cost” related to WTO dispute settlement not only consists of time, human resources and money spent on channeling a case through the increasingly complex stages of the system (time, resources and money which cannot therefore be spent on other matters). Other costs may be trade or broader political frictions or stalemates created by legalizing trade disputes with a binary win or lose outcome, or the fact that a compulsory and automatic DSU has made it more difficult to conclude new trade liberalizing agreements since the DSU’s creation in 1995. If so, the benefits of enhanced enforcement of (some parts of) the 1995 Uruguay Round agreements thanks to the DSU must be weighed against the costs of legalization including not seeing substantial new trade liberalization at the WTO since the WTO’s creation. Crucially, this cost-benefit analysis may be positive or in favor of the DSU for some agreements or some types of norms (say, national treatment) but negative for others (e.g. trade facilitation commitments where breach occurs less because of strategic defection but rather due to a lack of resources so that collaborative methods may be more cost-effective than legal enforcement backed up by trade retaliation).

4. Indicators for (good or bad) performance

Finally, the last step goes beyond the descriptive of measuring the effect and comparing the costs and benefits of WTO dispute settlement. Taking a more normative turn, the focus turns to explanations or indicators that may explain good or bad performance under one or more goals, for one or more agreements or rules. This can lead to both a diagnosis of existing problems and suggestions for reform. Shany lists both “structural indicators” and “processes employed by international courts” that may be outcome predictors for a court’s performance (Shany 2012). These are mostly internal material factors and include legal powers, personnel capacity, resources and structural independence. Relevant process features are access to justice, participation of stakeholders, transparency and costs.

Judicial independence as a driver of court effectiveness has been particularly controversial. Posner and Yo (2005) have argued that “independence prevents international tribunals from being effective” as they “can render decisions that conflict with the interests of state parties” and make states “reluctant to use international tribunals unless they have control over the judges”. Helfer and Slaughter (1997), in contrast, argue that the most effective tribunals are independent ones. The truth is probably somewhere in between and depends on the particular benchmark or court function looked at. Especially if one defines independent tribunals, as Posner and Yo do, as tribunals where “the judges are appointed in advance of any particular dispute and serve fixed terms » (think of the International Court of Justice, International Criminal Court or the WTO Appellate Body) -- as opposed to dependent tribunals which have judges that « are appointed by the state parties for the purpose of resolving a particular dispute »

such as the Iran-U.S. Claims Tribunal, Investor-State arbitration tribunals or GATT/WTO panellists – independent tribunals are more likely to perform well in terms of “norm-development effectiveness” and “regime legitimization”; dependent tribunals may perform better on the scale of resolving specific disputes between two parties (“case-specific effectiveness”). Since party-appointed adjudicators deciding one specific dispute (dependent tribunals) may be more closely attuned to the interests of the disputing parties, unburdened by broader systemic concerns, they may, indeed, be better situated to facilitate a judicial outcome palatable for the parties in dispute. However, case-specific effectiveness may then be achieved at the expense of other court functions or goals. All depends, once again, on goal identification.

Beyond independence, Helfer and Slaughter (1997) list another 12 material and social factors that, in their view, enhance the performance of supranational adjudication, distinguishing between factors within the control of states (such as the composition of the tribunal or functional or fact-finding capacity of the court), factors within the control of judges themselves (such as awareness of audience, incrementalism and quality of legal reasoning) and factors outside the control of either states or judges (such as relative cultural and political homogeneity of states subject to the court).

II. Assessment

Legal scholarship on WTO performance has focused on WTO dispute settlement. Too often, however, commentators have wrongly fixated on usage rates (500 requests for consultations in little more than 20 years; hard to say whether this is high or low without knowing the total number of WTO violations in place) or compliance with WTO panel or Appellate Body rulings (at around 90% very high, but saying little about causation or primary norm compliance). A better approach consists of (i) goal identification; (ii) measurement; (iii) establishing causation; (iv) assessing cost-effectiveness; and (v) considering explanations or indicators explaining good or bad performance. In this light, the WTO has performed better for some goals (resolving specific disputes, developing case law) as compared to others (primary norm compliance, promptly offering market access to traders). In addition, the jury remains out on whether these achievements are cost-effective or could have been reached by other means (e.g. transparency, peer review or more collaborative methods) at lower cost. Finally, one factor explaining the WTO’s good performance on some benchmarks is likely the “dependence” of WTO panels and the Appellate Body on WTO members, the fact that WTO adjudicators are jointly appointed by WTO members and the close interaction between WTO dispute settlement and WTO membership control.

F. Conclusion

The foregoing has focused on dimensions of the performance of the WTO as an organization from different disciplinary perspectives. Below we summarize key objectives that have emerged from our discussion that the WTO has been asked to perform (Table 1). What we observe is that the eye of the beholder problem is further accentuated as the baselines vary across disciplines. However, some objectives, such as the performance during crises and the role of the WTO as an ‘anchor’ for national trade policies is of interest to both economics and IR. These two fields also share an interest in the operation of specific ‘micro goals’ or functions of the WTO – an example is the impact of the various transparency mechanisms that have been established by the WTO membership and specific transparency-related functions that have been delegated to the Secretariat. Finally, all three disciplines have focused on the operation of the dispute settlement system – perhaps the only function of the WTO where this is clearly the case. This is an area where there has been a substantial amount of joint research undertaken by legal scholars and economists, as exemplified by the long-running project

initiated by Henrik Horn and Petros Mavroidis in 2001 to assess the WTO case law on an annual basis through a law and economics lens.¹¹

Two important take-away points can be distilled: First, the literature in all three fields is vibrant. There are a variety of objectives discussed and many conceptual and methodological tools offered to study WTO performance. International law mainly focuses on the degree of compliance with legal norms or dispute rulings, the way these can be applied domestically and how the case-law develops in ways that are consistent towards building a body of law. Research so far has not sufficiently engaged with social scientific ways to measure performance or effectiveness in more systematic ways. International relations puts most attention to how the institution (understood broadly) facilitates international cooperation and constrains powerful actors. There is an often implicit underlying assumption that for a system to work it has to contain unilateral action against trading partners and help to integrate weaker states into the system. Finally, not surprisingly the economic literature has focused most directly on the economic impacts of the WTO. While some of the economics literature has been focused on big conceptual questions about why trade agreements exist in the first place, there is an extensive literature that seeks to gauge both ‘macro’ goals and ‘micro’ objectives and functions.¹²

Table 1: Disciplinary objectives

Discipline	Key Objectives
IR	Rule of law against abuse of power
IR	Integrating emerging powers
IR	Performance in times of distress
IR	LDC Integration
IR	Enforcement/dispute settlement/compliance
Economics	Reducing trade barriers and agreeing to policy disciplines
Economics	Avoiding protectionism during economic crises
Economics	Increasing trade, integration and specialization
Economics	Enforcement of commitments
IL	Case-specific compliance
IL	Primary norm compliance
IL	Embeddedness effectiveness – direct effect
IL	Norm development

Note: Fields with grey shading are objectives that overlap across disciplines.

Second, different discursive tools have developed within the three disciplines. While similar phenomena are captured and assumptions shared, these concepts might work with very different connotations and labels. This at times makes cross-disciplinary debate less efficient and may at times prohibit it altogether. However, it also exposes the potential for cross-fertilization: it may support the legal field to further embrace empirical analysis, and assist economics to broaden its (sometimes) narrow focus on welfare analysis and related performance indicators. Finally, IR may serve more than in the past as a bridging function between IL and economics, bringing the politics more to the forefront of analyses of international cooperation.

¹¹ Papers are jointly authored by a lawyer and an economist and published each year in an issue of the *World Trade Review*. All papers can be downloaded from <http://globalgovernanceprogramme.eui.eu/wto-case-law-project/>.

¹² As mentioned, the extensive economics literature dealing with the effects of specific WTO agreements and disciplines could not be discussed in this paper due to space constraints.

To conclude, overall the WTO performance cannot be reduced to single indicators. A more holistic approach allows understanding the multitude of objectives an organization such as the WTO can serve and also sends a warning signal to having unrealistic expectations about the role of international institutions. At the same time sources for good or bad performance need to be exposed. Such an approach might in particular help policy-makers to better understand the limited effects of organizations they create.

References

- Anderson, K. 2014. "Contributions of the GATT/WTO to Global Economic Welfare: Empirical Evidence," ANU Working Paper 2014/15.
- Baracat, E., J. M. Finger, P. Thorne, L. Raul and J. Nogues 2013. *Sustaining Trade Reform: Institutional Lessons from Peru and Argentina*. Washington DC: World Bank.
- Bagwell, K. and R. Staiger 2002. *The Economics of the World Trading System*. Cambridge: MIT Press.
- Bagwell, K. and R. Staiger 2011. "What Do Trade Negotiators Negotiate About? Empirical Evidence from the WTO," *American Economic Review* 101(4):1238–73.
- Baccini, L. and S. Kim 2012. "Preventing Protectionism: International Institutions and Trade Policy," *Review of International Organizations* 7(4):369–98.
- Busch, M., E. Reinhardt, and G. Shaffer 2009. "Does Legal Capacity Matter? A Survey of WTO Members," *World Trade Review* 8(4):559–77.
- Chang, P. and M. Lee 2011. "The WTO Trade Effect," *Journal of International Economics* 85(1):53–71.
- Davey, W. 2009. "Compliance Problems in WTO Dispute Settlement," *Cornell International Law Journal* 42(1):119–28.
- Davis Christina 2012. *Why Adjudicate? Enforcing Trade Rules in the WTO*. Princeton: Princeton University Press
- De Bièvre, Dirk, Arlo Poletti and Aydin Yildirim, 2015, "About the melting of icebergs: Political and economic determinants of dispute initiation and resolution in the WTO," World Trade Forum conference paper.
- Downs, G., D. Rocke, and P. Barsoom 1996. "Is the Good News about Compliance Good News about Cooperation?" *International Organization* 50(3):379–406.
- Drezner, D. 2006. "The Viscosity of Global Governance: When is Forum-Shopping Expensive?," Paper presented at the 102nd Annual Meeting of the American Political Science Association, Philadelphia, PA
- Dutt, P., I. Mihov and T. van Zandt 2013. "The Impact of the WTO on the Intensive and the Extensive Margins of Trade," *Journal of International Economics*, 91(2):204–19.
- Elsig, M. 2010. "The World Trade Organization at Work: Performance in a Member-driven Milieu," *Review of International Organizations* 5(3):345–63.
- Elsig, M. 2016. WTO dispute settlement and the study of compliance.
- Elsig, M. and T. Cottier 2011. "Reforming the WTO: The Decision-making Triangle Revisited." In T. Cottier & M. Elsig (Eds.), *Governing the World Trade Organization: Past, Present and Beyond Doha*. Cambridge: Cambridge University Press.
- Elsig, M. and J. Eckhardt. 2015. "The Creation of the Multilateral Trade Court: Design and Experiential Learning," *World Trade Review* 14(S1):13–32.
- Elsig, M. and P. Stucki 2012. "Low-Income Developing Countries and WTO Litigation: Why Wake Up the Sleeping Dog?" *Review of International Political Economy* 19(2):292–316.
- Ethier, W. 2007. "The Theory of Trade Policy and Trade Agreements: A Critique," *European Journal of Political Economy* 23:605–23.

- Finger, J.M. 1979. "Trade Liberalization: A Public Choice Perspective." In R. Amacher, G. Haberler, and T. Willett (Eds.), *Challenges to a Liberal International Economic Order*. Washington DC: American Enterprise Institute.
- Finger, J.M. 2007. "Implementation and Imbalance: Dealing with the Hangover from the Uruguay Round," *Oxford Review of Economic Policy* 23(3):440–60.
- Finger, J.M. and J. Nogués 2006. *Safeguards and Antidumping in Latin American Trade Liberalization: Fighting Fire with Fire*. Washington DC: World Bank.
- Decreux, Y. and L. Fontagné 2015. "What Next for Multilateral Trade Talks? Quantifying the Role of Negotiation Modalities," *World Trade Review* 14(1): 29–43.
- Drezner, D. 2007. *All Politics Is Global: Explaining International Regulatory Regimes*. Princeton: Princeton University Press.
- Estevadeordal, A. C. Freund and E. Ornelas 2008. "Does Regionalism Affect Trade Liberalization toward Non-Members?," *Quarterly Journal of Economics* 123(4):1531-75.
- Evenett, S. and J. Fritz, 2015. *The Tide Turns? Trade, Protectionism, and Slowing Global Growth*. CEPR: Global Trade Alert.
- Francois, J. 1997. "External Bindings and the Credibility of Reform," In A. Galal, and B. Hoekman (Eds.), *Regional Partners in Global Markets*. London: Centre for Economic Policy Research.
- Francois, J. and W. Martin 2004. 'Commercial Policy, Bindings and Market Access' *European Economic Review* 48:665–79.
- Freund, C. and E. Ornelas 2010. "Regional Trade Agreements," *Annual Review of Economics* 2(1):139–66.
- Fronk, J. 2014. "Do Procurement Agreements Work?" Georgetown University. At: <http://jaredfronk.wix.com/jobmarket>
- Gawande, K., B. Hoekman and Y. Cui 2015. "Global Supply Chains and Trade Policy Responses to the 2008 Financial Crisis," *World Bank Economic Review* 29(1):102–28.
- Gutner, T. and A. Thompson 2010. "The Politics of IO Performance: a Framework," *Review of International Organizations* 5(3)
- Grant, R. and R. Keohane 2005. "Accountability and Abuses of Power in World Politics," *American Political Science Review* 99(1):29–43.
- Gruber L. 2000. *Ruling the World: Power Politics and the Rise of Supranational Institutions*. Princeton: Princeton University Press.
- Handley, K. 2014. "Exporting under Trade Policy Uncertainty: Theory and Evidence," *Journal of International Economics* 94(1):50–66.
- Helfer, L. 2014. "The Effectiveness of International Adjudicators." In K. Alter, C. Romano & Y. Shany (Eds). *Oxford Handbook of International Adjudication*, Oxford: Oxford University Press, pp. 464-482.
- Helfer, L. and A-M. Slaughter 1997. "Toward a Theory of Effective Supranational Adjudication". 107 *Yale L.J.* 273.
- Hoekman, B. (Ed.) 2015. *The Global Trade Slowdown: A New Normal?* London: CEPR Press.
- Hoekman, B. and A. Nicita 2011. "Trade Policy, Trade Costs and Developing Country Trade," *World Development*, 39(12): 2069-79
- Horn, H., Nordström, H. and P.C Mavroidis 1999. "Is the use of the WTO dispute settlement system biased?" CEPR Discussion Paper 2340.

- Hughes, V. 2012. "Why Is the WTO Dispute Settlement System So Successful?" Video lecture available at <http://mdsweb.vuw.ac.nz/Mediasite/Viewer/Viewers/Viewer320TL.aspx?mode=Default&peid=375b3db3-0b6d-4d8e-b3b0-7d45a9d2d904&pid=3daf095e-e1a8-4f1f-8ec2-0abfd39640be&playerType=SL1>
- Johnson, H. 1953-54. "Optimum Tariffs and Retaliation," *Review of Economic Studies* 21(2):142–53.
- Keohane R. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton: Princeton University Press.
- Kim, M. 2008. "Costly Procedures: Divergent Effects of Legalization in the GATT/WTO Dispute Settlement Procedures," *International Studies Quarterly* 52(3):657–86.
- Laborde, D and W. Martin 2015. "Formulas for Failure? Were the Doha Tariff Formulas too Ambitious for Success?," *World Trade Review* 14(1):45-65.
- Lang, A. and J. Scott 2009. "The Hidden World of WTO Governance," *European Journal of International Law* 20(3):575–614.
- Limão, N. and P. Tovar 2011. "Policy choice: Theory and Evidence from Commitment via International Trade Agreements," *Journal of International Economics* 85(2):186–205.
- Liu, X. 2009. "GATT/WTO Promotes Trade Strongly: Sample Selection and Model Specification," *Review of International Economics* 17(3):428–46.
- Maggi, G. and A. Rodriguez-Clare 1998. "The Value of Trade Agreements in the Presence of Political Pressures," *Journal of Political Economy* 106(3):574–601.
- Martin, W. and P. Messerlin 2007. "Why Is It So Difficult? Trade Liberalization Under the Doha Agenda," *Oxford Review of Economic Policy* 23(3):347–66.
- Pauwelyn, J. 2005. "The Transformation of World Trade," *Michigan Law Review* 104:1-70.
- Pelc, K. 2010. "Constraining Coercion? Legitimacy and Its Role in US Trade Policy, 1975–2000," *International Organization* 64(1): 65–96.
- Posner, E. and J. Yoo 2005. "Judicial Independence in International Tribunals," *California Law Review* 93
- Raustiala, K. 2000. "Compliance and Effectiveness in International Regulatory Cooperation," *Case W. Res. J. Int'l L.* 32:387
- Rickard, S. and D. Kono, 2014. "Think Globally, Buy Locally: International Agreements and Government Procurement," *Review of International Organizations* 9(3):333–52.
- Rose, A. 2004. "Do We Really Know That the WTO Increases Trade?," *American Economic Review* 94(1):98–114.
- Rose, A. 2005. "Which International Institutions Promote International Trade?" *Review of International Economics* 13(4): 682–98.
- Sattler, T. and T. Bernauer 2011. "Gravitation or Discrimination? Determinants of Litigation in the World Trade Organization," *European Journal of Political Research* 50(2):143–67.
- Shany, Y. 2012. "Assessing the Effectiveness of International Courts: A Goal-Based Approach," *American Journal of International Law* 106(2):225–70
- Shingal, A. 2011. "Services Procurement under the WTO's Agreement on Government Procurement: Whither Market Access? » *World Trade Review* 10(4):527–49.

- Shingal, A. 2015. "Econometric Analyses of Home Bias in Government Procurement," *Review of International Economics* 23:188–219.
- Subramanian, A. and S. Wei 2007. "The WTO Promotes Trade, Strongly but Unevenly," *Journal of International Economics* 72(1):151–175
- Tang, M. and S. Wei 2009. "The Value of Making Commitments Externally: Evidence from WTO Accessions," *Journal of International Economics* 78(2): 216–29.
- Tumlir, J. 1985. *Protectionism: Trade Policy in Democratic Societies*. Washington, DC: American Enterprise Institute.
- Wilson, B. 2007. "Compliance by WTO Members with Adverse WTO Dispute Settlement Rulings: The Record to Date," *Journal of International Economic Law* 10(2):397–403
- WTO, 2015. "WTO disputes reach 500 mark", News Item, 10 November 2015, available at https://www.wto.org/english/news_e/news15_e/ds500rfc_10nov15_e.htm
- Zangl, B. 2008. "Judicialization Matters! A Comparison of Dispute Settlement Under GATT and the WTO," *International Studies Quarterly* 52(4):825–54.

Author contacts:

Manfred Elsig

University of Bern and World Trade Institute

Email: Manfred.Elsig@wti.org

Bernard Hoekman

Robert Schuman Centre for Advanced Studies, EUI

Villa La Fonte

Via delle Fontanelle, 18

50014 San Domenico di Fiesole (FI)

Italy

Email: bernard.hoekman@eui.eu

Joost Pauwelyn

Graduate Institute of International and Development Studies

Email: joost.pauwelyn@graduateinstitute.ch