

How Effective Are Regional Trade Agreements?  
Ask the Experts

Julia Gray  
Assistant Professor  
University of Pittsburgh

Jonathan B. Slapin  
Assistant Professor  
University of Houston

September 30, 2010

## **Abstract**

Limited data on international organizations prevents us from testing some of that literature's most important claims. Expert surveys — a technique that is widely used in comparative and American politics but thus far all but neglected in international relations — can allow us to capture many aspects of international organizations for which data would otherwise be unavailable. They also provide us with a means of comparing agreements across regions and issue areas. We present a new dataset of expert opinion on the world's regional trade organizations. Our survey covers 25 dimensions of 45 different regional agreements — including the effectiveness of trade liberalization, the effectiveness of dispute settlement, and the political and international influence of the agreement. Bayesian factor analysis allows us to extract a principle dimension of RTA effectiveness.

# 1 Introduction

Regional trade agreements of all variety span the globe. As more and more states form these associations, they have become an ever more popular area of academic enquiry. Yet conducting empirical research involving regional trade agreements is problematic. Social scientists know that these agreements differ widely, and that the Arab Maghrebi Union probably works very differently in practice than the European Union. Yet how can we capture those differences? Many researchers have turned to the elements of endogenous design of the agreement (Keohane, Moravcsik and Slaughter 2000; McCall Smith 2000) as a proxy for the effectiveness of an agreement. In other words, they examine what states have chosen to write into the agreement. But legal charters are easy to copy and in many cases just as easy to ignore. Others look at the observable economic outcomes for states in an RTA — such as increased trade or drops in tariff. These outcomes, though, do not necessarily capture the independent effects of a particular agreement. That is, since the unit of analysis in that strategy is member states rather than the agreement itself, we do not know how much of those economic changes to attribute to the agreement itself. Furthermore, it is difficult to map whether those economic outputs are on target with the initial goals of the organization. Lastly, many RTAs cover several different areas of competence and may not excel at all of them. How, then, can we meaningfully measure and compare the workings of regional agreements?

This paper uses expert surveys — a method that is widely used in comparative and American politics but has yet to be applied to the study of international organizations — to explore the characteristics of regional trade agreements. Expert surveys offer considerable leverage in environments where hard data for the types of things we want to measure simply do not exist. By talking directly to the individuals who have worked with these organizations — either as negotiators, consultants, or actual staff members — we are able to get a clear picture of the actual functioning of these agreements on many dimensions.

This is a significant contribution to the empirical study of international organizations,

both in terms of the method and the insights derived from our survey. Much of the literature in international organizations hinges on concepts that are difficult to measure. Though survey methods are gaining leverage in IO research more generally (Tomz 2007), expert surveys have yet to be fully exploited, and they offer great potential for capturing dimensions of IO activity that may otherwise be challenging to quantify. They also allow us to compare organizations in fine detail and across like units. This is a significant improvement from the current state of affairs, where researchers must either use relatively un-nuanced measurements such as dummy variables to compare the effects of organizations across the world (Milner and Büthe 2008; Mansfield and Reinhardt 2008), or where they focus on one organization in great detail to explore its workings (Moravcsik 1998; Chase 2003; Alter, Helfer and Guertzovich 2009). Survey data allow social scientists to quantify many organizations in like units. They can be used in subsequent analysis either as dependent variables (that is, what types of factors tend to be associated with the creation of organizations with particular characteristics?) or as independent variables (how do certain IO characteristics impact a set of outcomes?).

For the purposes of this paper, we present the results of surveys of experts around the globe to show how one subset of international organizations — that is, regional trade agreements — works in practice. We gathered data on 25 different characteristics for RTAs encompassing their potential capacity, including but not limited to how well they work as trade promoting instruments, how they deal with non-tariff barriers, how well their bureaucracies function, their perceived influence, the effectiveness of adjudication, and the match or mismatch between their ambitions and their actual competencies. The output, in fact, is striking. We use a Bayesian factor analysis model to extract a single principal dimension of RTA effectiveness from these expert responses. The RTAs fall along this dimension as we would expect, with those RTAs generally perceived to work well at one end, and those seen as less effective at other end. That is, despite their having many different competencies, agreements simply tend to work or they do not, with little nuance in between.

This paper proceeds as follows. The next section provides a theoretical foundation by sketching out some of the gaps in our understanding of regional trade agreements and describes in greater detail how expert surveys can fill those gaps. The subsequent section presents our survey instrument and methodology. The following section is an empirical analysis of the data we gathered. We first display summary statistics of both the dimensions gathered and how the RTAs in question score on those dimensions. We then use Bayesian factor analysis to identify a single common factor underlying the expert scores, and show how the RTAs in our survey rank along this dimension. The final section concludes.

## **2 Toward A Better Understanding of Regional Integration**

The study of all types of international organizations is plagued by a lack of good comparative data, and many researchers have already bemoaned this reality (Kahler, Burton and Montgomery 2009; Kratochwil and Ruggie 1986; Oneal and Russett 2005; King 2001; Mastenbroek 2005). Many of the big questions about international organizations center on whether IOs actually have any independent effect on their member states' behavior. For many researchers, this hinges on their ability to enforce their own rules (Drezner 1999), the level of legal obligation imposed on their members (Busch and Pelc 2010; Simmons 2000*a,b*), the quality and efficiency of their staff (Vaubel, Dreher and Soylu 2007) — indeed, their overall effectiveness.

Regional trade agreements are no exception. Some academic work focused on the economic aspects of these agreements and their effect on multilateralism as a whole (Bhagwati, Greenaway and Panagariya 1998), or the phenomenon of regionalism more generally (Mansfield and Milner 1999; Pomfret 2007; Whalley 2008). Not all work, though, tackled head-on the variation across these agreements, with exceptions in economics (Grossman and Helpman 1995) as well as political science (Mansfield, Milner and Pevehouse 2008;

Haftel 2007). Yet substantial variation exists across these agreements, and it might initially seem that this variation would be easy to measure. Because they are organizations that are primarily concerned with economic output — that is, the enhancement of trade volumes and the decrease of trade barriers — at first glance it seems as though they might be easier to quantify than organizations who have more amorphous goals, such as the promotion of cooperation more generally or a decrease in conflict. Trade volumes as well as tariff levels can be measured, of course, on a country-by-country, annual basis. Similarly, trade agreements vary in terms of their levels of proposed integration. This is often quantified as six possible stages, from a preferential area to a free trade area to a customs union to economic and monetary union to complete economic integration.

However, although these indicators seem at first glance to be straightforward ways of capturing not just variation across RTAs but also their effectiveness, at best their usefulness may be limited, and at worst they may be deceptive. For example, many have noted that as tariffs fall, they are simply displaced by nontariff barriers (NTBs) such as quotas, countervailing duties, and voluntary export restraints. The net effect could be that the overall level of trade protection may stay the same or even increase (Anderson and Schmitt 2003; Mansfield and Busch 1995). This would indicate that focusing on reduction of tariff barriers alone could distort our perception of the effectiveness of a given RTA, if NTBs — which are notoriously difficult to measure — rose apace.

Looking at changes in the aggregate levels of trade can also be misleading. Trade volumes can only be gathered on a member-state basis — that is, at the country level. How do we know whether to attribute any changes in trade levels to a particular RTA? It could well be the case that liberalization would have been achieved anyway and that states simply sign RTAs that accord with their preexisting intentions.<sup>1</sup> If this were the case, we would still not be any closer to identifying the independent power of RTAs even if we did see an increase in trade volumes for their member states. Furthermore, many countries are members of multiple agreements, both at the multilateral, bilateral, and

---

<sup>1</sup>This is the well known problem of selection bias; see Heckman (1976); Vreeland (2003); VonStein (2005) for applications.

regional level. If trade increases for a given country, how do we know which agreement is doing the work?

Finally, an examination of the levels of proposed integration may fall short of what is achieved in reality. Many economic agreements propose levels of economic coordination that they are slow to achieve — if, indeed, they achieve them at all. Bearce and Omori (Bearce and Omori 2005), when coding the level of actual integration achieved by many agreements, found a significant gap between what was written on paper and what played out in practice. Although this phenomenon is widespread, the South African Development Community (SADC) represents just one example. Though it has faltered in even establishing the basics of free trade agreement, it continues to set targets for deeper integration whose deadlines continuously are broken. Its plans to establish a common market by 2010 are set to fall through, as most observers agree that it barely even has a functioning free trade area. “There is now recognition that you cannot talk about a common external tariff if you have no common policies,” the deputy director of South Africa’s Department of Trade and Industry was quoted as saying.<sup>2</sup> Thus, *de jure* integration may not match *de facto* in many cases.

Even if we go back to the classic work on international organizations to look at the legal language or dispute settlement capacity of an RTA as a proxy for overall effectiveness, that may not tell the full story. There can be a substantial difference between the legal language of a given agreement — which is easy enough to write or even copy — and the capacity of a court in practice (Gray and Slapin 2009). These differences may arise for a variety of reasons, including domination by powerful states of the adjudication process, a lack of recognition of the authority of the court, or the inability for decisions to be binding in practice at the legal level of the member states in practice. For example, though the Common Market of the South (Mercosur) has a tribunal that looks good on paper, it is widely acknowledged that the largest members of the agreement — Brazil and Argentina — dominate the process. “Disputes are arbitrated according to the relative weight of

---

<sup>2</sup>“SADC countries fail to meet customs union, currency targets,” PANA, 4 September 2009.

the member countries,” one commentator observed. Indeed, Paraguayan representative Wilfrido Fernández resigned from the Tribunal in October 2008, on the grounds that member states were “not showing the political will” for the court to work effectively.<sup>3</sup>

In an effort to better quantify the many aspects of regional trade organizations, some researchers have embarked on efforts to code the breadth (the number of tariff lines covered), the depth (the extent of reduction and elimination in barriers) and the rate (the phase-in time of reduction and elimination) of trade liberalization (Kim and Hicks 2008; Kim 2010; Baccini and Dur 2009).<sup>4</sup> These commendable efforts go a long way toward deepening our understanding of the many dimensions of regional organizations. A rigorous, detailed and systematic approach to coding integration agreements can only be beneficial in capturing how these agreements work.

Still, though, we are left with a few problems. First, these aforementioned coding endeavors are often so encompassing and so specific that they end up focusing on one region or country. Kim 2010, for example, codes US PTAs; Baccini and Dur 2009 focus on the European Union’s Economic Partnership Arrangements; and Kim and Hicks (Kim and Hicks 2008) quantify various aspects of Asian PTAs. Part of this has to do with the enormity of the task, but another part may well be that each country or region has a type of agreement that is specific to them. Thus, we are left again with metrics that are not truly comparative across regions.

Second, as mentioned above, particularly with regional trade agreements, there seems to be a broad acknowledgement that they often do not work in well in practice as they do in theory. Thus, even these metrics may fail to capture the implicit understanding that many agreements are not what they seem. This dimension is certainly knowable but is difficult to get at with existing data or even through coding exercises. How can we capture, then, this latent understanding of the workings of regional trade agreement in a way that is comparable across agreements in different parts of the world?

---

<sup>3</sup>“Mercosur Limpes Slowly Along,” Mario Osava, Interpress Service, 3 January 2008.

<sup>4</sup>Indeed, many other researchers have embarked on similar efforts with IGOs more generally; see Koremenos ongoing, Boehmer, Gartzke and Nordstrom (2004); Volgy et al. (2008); Bondanella (2009).



### 3 Expert Surveys and their Application to Regional Agreements

To tackle this problem, we turn to expert surveys. The reliance on experts for judgments, and the use of those judgements as an independent source of data for analysis, is widespread in political science — particularly in comparative politics — as a means of acquiring data where none exist, particularly perceptions-based data (Laver 1998; Whitefield et al. 2007; Hooghe, Marks and Wilson 2002; Benoit and Laver 2006; Marks et al. 2007; Busch, Reinhardt and Schaffer 2009). However, this method is — regrettably, we argue — infrequently used in international relations generally and in the study of international organizations more specifically. Expert surveys are widely acknowledged as a valuable way of obtaining high-level opinions on political and other phenomena, from party positions to the functioning of organizations such as the European Commission.

The goal of expert surveys differs from that of much survey research, which usually focuses on trying to link characteristics of particular groups with particular outcomes. By contrast, because of the high level of many of these respondents, very little personal data on individual characteristics is typically culled — in fact, most experts typically will only consent to be surveyed on the grounds of anonymity, since they do not want to compromise their own reputation or that of their affiliated institution. Thus, the idea here is usually to use the consensus of these experts as a way of generating data on dimensions that are otherwise difficult to capture.

Of course, there are drawbacks to this approach. The main one is the relatively small number of observations that expert surveys often yield. This stands in stark contrast to survey research in, say, studies of political opinion, where tens of thousands of survey respondents are often culled — or even in most time-series cross-section datasets. However, this is an inevitable function of the universe of cases. Combine the relatively low response rates of many surveys — the standard is between 20 and 30 percent — with the much smaller pool of potential respondents once we move to the expert levels, and

small Ns are an unavoidable consequence. Another is respondent bias, which might be of particular concern with a limited number of respondents and where the bias of one particular individual may skew results. An expert may have a particular stance based on his or her country of origin, organizational affiliation, or political or personal leanings. Though this bias is unavoidable, it is at least in part detectable and can be discounted by accounting for their workplace as well as their region of origin.

In sum, expert surveys offer significant advantages for the study of international organizations more generally and regional organizations in particular. Such surveys enable us not only to capture dimensions of international organization that are otherwise difficult to measure directly, but also to develop a rubric that makes those dimensions comparable across units. The following section describes the survey instrument that we developed and its implementation.

## **4 Data Gathering Method and Instrument**

We based our survey instrument on Benoit and Laver's party positions instrument, which allows experts to establish overall rankings of subjects on their performance on a particular issue as well of the salience of that issue to the subject in question. Koremenos et al 2001 call for IOs to be coded membership rules, scope of issues covered, centralization of tasks, rules for controlling the institution and flexibility of arrangements. To that end, we included 25 questions covering exactly those areas. The full list of questions, in the order in which they were asked, appear in the appendix. Experts were asked to fill out as many dimensions for as many RTAs that they felt qualified to judge, leading to a total of 3,831 individual data points that cover various aspects of 35 different trade agreements.

These surveys were accompanied with an explanatory cover letter which described the goal of the project and provided instructions for completing the survey, as well as an example. The 25 questions were presented in a manner similar to feeling thermometers and had two parts: the substantive dimension as well as its salience. Each question

had a general label and then was accompanied by descriptors identifying the highest and lowest points of sentiment, on a scale of one to ten. For example, for the general question on *Political Influence*, the lowest point (1) was identified as “No political influence on members and region,” compared with a 10, which we identified as “High degree of political influence on members and region.” This was the first part of the question; the second part, for all questions, was the salience of that particular issue — that is, how important a particular dimension was for the RTA in question, also scaled from one to ten. This is an important component, since not all 25 issues are equally important for each RTA, depending on its mission statement and ambitions. Political influence, for example, may not be a stated goal for NAFTA, so even though it might receive a low score for political influence, its salience would also be low-ranked for that particular issue. An organization such as the South African Development Community, however, has ambitions for political influence but (many believe) falls short of those ambitions. Thus, it might receive a similarly low score for political influence, but the salience would be highly ranked; it is an issue that SADC itself at least considers important but fails to achieve. Therefore, including these salience dimensions helps us to understand the occasional gap between an RTA’s stated goals and their performance on those goals.

We conducted these surveys between 2008 and summer 2009. We attempted to distribute these surveys electronically in the initial stages but met with little response, so almost all surveys were conducted in person. Our sample of respondents includes representatives from trade negotiating offices, think tanks in Europe, southern Africa, and Latin America, for a total of 25 experts. This number might seem small, but it is important to remember that response rates for surveys of all stripes are usually around 20 percent — and the number of individuals that are both qualified to judge the workings of regional agreements is a very limited set. In the Chapel Hill expert survey on party positions in Europe, the researchers obtained 98 survey responses (a 34 percent response rate) (Steenbergen and Marks 2007; Marks et al. 2007). Unlike regional trade agreements, parties are far more public and a far greater number of individuals are qualified

to comment on their workings.

The following section describes the data we gathered and analyzes these dimensions across RTAs.

## 5 Describing the Data

Tables One through Three show descriptive statistics for the surveys we gathered. The first table shows summary statistics for every RTA that received a response. Tables Two and Three show the summaries for the data after we collapsed them by overall scores for each RTA. Thus, the Ns represent not the number of respondents but the number of RTAs that received scores on that particular dimension.

TABLES 1-3 ABOUT HERE

The first thing to notice in Table One, which shows the scores assigned to each RTA across the relevant dimensions, is that the standard deviations for each score are relatively small. The average SD, in fact, is 2.8, indicating that respondents deviated from their responses by less than 3 points on our ten-point scale. This indicates a relatively high degree of expert consensus about the actual workings of the world's RTAs.

In Tables Two and Three, it is interesting to note the degree of variance *across* RTAs. Note that although experts made use of the full range of the scale — the minimum and maximum scores given for both the means and the salience almost always run the complete spectrum — there is substantial variance in RTA performance, resulting in average scores across all RTAs somewhere in the middle of the spectrum. The dimension on which RTAs seem, on average, to perform the best are WTO compliance (mean score 7.13 out of a possible 10) and ambition (6.89). Note, however, that their ability to achieve those ambitious goals is somewhat more limited (4.71). Progress on trade in services is the area in which RTAs do the worst on average (3.47 out of 10), with an RTA's enforcement of its own procedures not far behind (3.72). Similarly, the actual performance of dispute settlement mechanisms is relatively low (4.18) — in fact, this dimension is the fifth-worst

performing one on average across RTAs, despite the hype given to this feature of RTAs in much of the IR and IPE literature.

How related, though, are these dimensions? That is, how linked is RTA success in one area to that in another? Table Four shows a correlation matrix across all the dimensions surveyed.

#### TABLE FOUR ABOUT HERE

If we take each RTA’s average score for “ability to meet its own goals” as an overall proxy for competence, we see that this correlates most closely with a well-functioning dispute settlement mechanism, high levels of legalization, and a well-functioning secretariat ( $r=.99$  for all three dimensions). Closely behind are the dimensions that represent effectiveness in trade — correlations with ability to trade in goods and services and WTO compliance are above  $.90$ . This indicates that well-functioning bureaucracies and courts also tend to go hand in hand with effectiveness in trade promotion.

These statistics, however, are a relatively basic way of mapping the similarities across these dimensions. For a more detailed look, we turn to factor analysis, as described in the section below.

### 5.1 Bayesian Factor Analysis

We next turn to a Bayesian factor analysis model to extract a principal dimension of effectiveness from these dimensions. Factor analysis allows us to empirically capture latent constructs underlying a particular set of variables. Bayesian factor analysis particularly useful for two reasons. First, it provides a measure of uncertainty around the estimated factor scores, and second, in a Bayesian framework it is still possible to estimate the model even when not every RTA is scored on every dimension. This is important in our context as our experts were unable to score every RTA on every dimension.

Since we are interested in the primary dimension of integration, we run a one-dimensional factor model. Because of the relatively high degree of missingness — experts often did

not comment on every single dimension for each RTA that they chose to address — we limited the analysis to RTAs and dimensions that had at least 20 observations.

We collapsed the mean and the salience variables into one, by first rescaling the main scores from 1 to 10, to -5 to 5, before multiplying that rescaled score by the salience of that particular dimension. This allows us to capture the score of a particular dimension along with its relative importance. For example, if an agreement scores below average on a particular dimension but the salience for that dimension to that agreement is above average, its score will be negative, indicating a failure to live up to its own expectations.

Tables Five and Six show the results of this analysis for the dimensions of RTA competence in our dataset. The  $\lambda$  scores represent the factor loading matrix — that is, the degree to which any given dimension maps on to the single common factor. The  $\psi$  scores represent uniqueness, or the singularity of any dimension’s contribution to the factor. Lastly, the factor scores  $\phi$  place the RTAs onto this dimension. These are the score that we are most interested in.

TABLES FIVE AND SIX ABOUT HERE

FIGURE 1 ABOUT HERE

Figure 1 presents the factor scores for all the RTAs, ranking them from lowest to highest. In addition, we plot the 95% Bayesian credible intervals around these scores. The European Union is by far the best-performing RTA in the dataset, with a significant gap below even the second-best RTA. However, the confidence interval around the EU’s score is quite large, meaning we cannot rule out with 95% confidence that it has a higher score than some other well performing agreements such as EFTA, NAFTA, CEFTA, and OECS. Other agreements have a middling level of effectiveness, such as the Andean Community. Lastly, the organizations deemed to be the worst performing include Mercosur, Asean, Sparteca, and APEC.

If we think of the latent first dimension as being organized around effective institutions that promote trade, the bottom end of that scale makes some sense. APEC, despite being

an economic agreement and having trade liberalization as one of its pillars, does not actually constitute a free trade area, and ASEAN's was established relatively recently in the organization's history. Similarly, despite relatively low levels of institutionalization, NAFTA, CEFTA and EFTA have done relatively well in promoting trade. What is interesting is that this trade dimension seems to lie beneath many of the other dimensions capturing effectiveness; that is, that if a RTA does not succeed in promoting trade, it will fall short in other areas of competence as well.

## 6 Conclusion

This paper makes several contributions to the study of international organizations. First, we have argued for the more frequent application of expert surveys to capture the workings of international agreements in a comparative manner. This method is underused in IO research at present and offers significant potential for investigating many of the claims about the workings of international organizations. Through direct surveying of the experts who have regular contact with these organizations, we can gain valuable insight on attributes of these organizations that are otherwise difficult to measure.

Second, we have shown that, for regional trade agreements, there tends to be one underlying factor of competence across which these agreements can be sorted. Bayesian factor analysis showed that RTAs can effectively be sorted by a single dimension of competence. This indicates that, despite the many areas that RTAs can potentially cover and the differences in their goals and scope, RTAs tend to either work well across all dimensions, or they do not.

This is an important step in our understanding of how regional trade agreements work. Further analysis and ongoing research using these data can explore the domestic or international determinants of RTA competence (that is, using the survey data as a dependent variable), as well as the influence that these dimensions might have on other outcomes of interest (survey data as independent variables). For now, we hope to have

established a framework for future analysis of how regional agreements work in practice, and to have established a way of making meaningful comparisons across agreements.



## References

- Alter, Karen, Laurence Helfer and Maria Florencia Guerzovich. 2009. "Islands of Effective International Adjudication: Constructing an Intellectual Property Rule of Law in the Andean Community." *American Journal of International Law* 103(1):1–47.
- Anderson, Simon P. and Nicolas Schmitt. 2003. "Nontariff Barriers and Trade Liberalization." *Economic Inquiry* 41(1):80–97.
- Baccini, Leo and A. Dur. 2009. "The New Regionalism and Policy Interdependence." Manuscript, University College Dublin.
- Bearce, David H. and Sawa Omori. 2005. "How Do Commercial Institutions Promote Peace?" *Journal of Peace Research* 42(6):659–678.
- Benoit, Kenneth and Michael Laver. 2006. *Party Policy in Modern Democracies*. London: Routledge.
- Bhagwati, Jagdish, David Greenaway and Arvind Panagariya. 1998. "Trading Preferentially: Theory and Policy." *The Economic Journal* 108(449):1128–1148.
- Boehmer, Charles, Erik Gartzke and Timothy Nordstrom. 2004. "Do Intergovernmental Organizations Promote Peace?" *World Politics* 57(1):1–38.
- Bondanella, Stacy. 2009. Intergovernmental Organizations and the Determinants of Member State Interest Convergence PhD thesis University of Pittsburgh.
- Busch, Marc and Krzysztof Pelc. 2010. "The Politics of Judicial Economy at the World Trade Organization." *International Organization* 64(2):257–280.
- Busch, Marc L., Eric Reinhardt and Gregory Schaffer. 2009. "Does Legal Capacity Matter? A Survey of WTO Members." *World Trade Review* 8(4):559–577.
- Chase, Kerry A. 2003. "Economic Interests and Regional Trading Arrangements: The Case of NAFTA." *International Organization* 57(1):137–174.

- Drezner, Daniel. 1999. *The Sanctions Paradox: Economic Statecraft and International Relations*. Cambridge University Press.
- Gray, Julia and Jonathan Slapin. 2009. "Why Some Regional Trade Agreements Work: Ambition, Outside Options, and Legalization." Midwest Political Science Association 67th Annual National Conference.
- Grossman, Gene M. and Alhanan Helpman. 1995. "The Politics of Free-Trade Agreements." *American Economic Review* 85(4):667–690.
- Haftel, Yoram Z. 2007. "Designing for Peace: Regional Integration Arrangements, Institutional Variation and Militarized Interstate Disputes." *International Organization* 61(1):217–237.
- Heckman, James. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5(4):475–492.
- Hooghe, L., Gary Marks and C Wilson. 2002. "Does left/right structure party positions on European integration?" *Comparative Political Studies* 35:965–989.
- Kahler, Miles, Emilie Hafner Burton and Alexander H. Montgomery. 2009. "Network Analysis for International Relations." *International ORganization* 63:559–592.
- Keohane, Robert O., Andrew Moravcsik and Anne-Marie Slaughter. 2000. "Legalized Dispute Resolution: Interstate and Transnational." *International Organization* 54(3):457–488.
- Kim, Moonhawk. 2010. "Bringing Trade Liberalization Back into Trade Agreements: Explaining the Variation in U.S. Preferential Trade Agreements." Working paper.
- Kim, Soo Yeon and Raymond Hicks. 2008. "Commitment, Signaling, or Flexibility? The Effectiveness of PTAs in the Asia-Pacific." Paper presented at the annual meeting of

- the International Political Economy Society, University of Pennsylvania, Philadelphia, PA, November 12, 2008.
- King, G. 2001. "Proper nouns and methodological propriety: Pooling dyads in international relations data." *International Organization* 55(2):497–507.
- Koremenos, Barbara, Charles Lipson and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55(4):761–700.
- Kratochwil, F. and J.G. Ruggie. 1986. "A State of the Art on an Art of the State." *International Organization* 40:753–775.
- Laver, Michael. 1998. "Party Policy in Britain 1997: Results from an Expert Survey." *Political Studies* 46(2).
- Mansfield, E.D., Helen V. Milner and Jon C. Pevehouse. 2008. "Democracy, Veto Players and the depth of Regional Integration." *The World Economy* 31(1):67–96.
- Mansfield, Edward D. and Eric Reinhardt. 2008. "International Institutions and the Volatility of International Trade." *International Organization* 62(4):621–652.
- Mansfield, Edward D. and Helen V. Milner. 1999. "The New Wave of Regionalism." *International Organization* 53(3):589–627.
- Mansfield, Edward D. and Marc L. Busch. 1995. "The Political Economy of Nontariff Barriers: A Cross-National Analysis." *International Organization* 49(4):723–749.
- Marks, G., L. Hooghe, M.R. Steenbergen and R. Bakker. 2007. "Crossvalidating data on party positioning on European integration." *Electoral Studies* 26(1):23–38.
- Mastenbroek, E. 2005. "EU compliance: Still a 'black hole'?" *Journal of European Public Policy* 12(6):103–1120.
- McCall Smith, James. 2000. "The Politics of Dispute Settlement Design: Explaining Legalism in Regional Trade Pacts." *International Organization* 54(1):137–180.

- Milner, Helen and Tim Büthe. 2008. "The Politics of Foreign Direct Investment into Developing Countries: Increasing FDI through International Trade Agreements." *American Journal of Political Science* 52(4):741–762.
- Moravcsik, Andrew. 1998. *The Choice For Europe*. Ithaca, NY: Cornell University Press.
- Oneal, J.R. and B. Russett. 2005. "Rule of three, let it be? When more really is better." *Conflict management and Peace Science* 22:293–310.
- Pomfret, Richard. 2007. "Is Regionalism an Increasing Feature of the World Economy?" *Journal of Economic Surveys* 30(6):923–947.
- Simmons, Beth. 2000a. "International Law and State Behavior: Commitment and Compliance in International Monetary Affairs." *American Political Science Review* 94(4):819–835.
- Simmons, Beth. 2000b. "The Legalization of International Monetary Affairs." *International Organization* 54(3):573–602.
- Steenbergen, Marco R. and Gary Marks. 2007. "Evaluating Expert Surveys." *European Journal of Political Research* 46(3):347–366.
- Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61(4):821–840.
- Vaubel, Roland, Axel Dreher and Ugurlu Soylu. 2007. "Staff Growth in International Organizations: A Principal-Agent Problem? An Empirical Analysis." *Public Choice* 133:275–295.
- Volgy, Thomas J., Keith A. Grant, Elizabeth Fausett and Stuart Rodgers. 2008. "Identifying Formal Intergovernmental Organizations." *Journal of Peace Research* 45(6):837–850.
- VonStein, Jana. 2005. "Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance." *American Political Science Review* 99:611–622.

- Vreeland, James Raymond. 2003. *The IMF and Economic Development*. New York: Cambridge University Press.
- Whalley, Jon. 2008. "Recent Regional Agreements: Why So Many, Why So Much Variance in Form, Why Coming So Fast, and Where Are They Headed?" *World Economy* 31(4):517–532.
- Whitefield, Stephen, Milada Anna Vachudova, Marco R. Steenbergen, Robert Rohrschneider, Gary Marks, Matthew P. Loveless and Liesbet Hooghe. 2007. "Do expert surveys produce consistent estimates of party stances on European integration? Comparing expert surveys in the difficult case of Central and Eastern Europe." *Electoral Studies* 26:50–61.

## 7 Appendix

### Questions on Expert Survey (Scale of 1-10)

1. Trade Negotiator Capacity (in aggregate): Trade negotiators are incompetent and ineffective/Trade negotiators are professionalized, competent and well-informed
2. Quality of Infrastructure/Secretariat: Has no staff and no separate bureaucratic structure/Has a well-trained staff and well-organized bureaucracy
3. International Influence : The organization as an entity has no influence on international affairs /The organization as an entity has significant influence on international affairs
4. Trade Diversion: Has effectively increased trade with third parties (trade creation)/Has effectively decreased trade with third parties (high trade diversion)
5. Market Access in Services: Provides members no market access in services/Provides members full market access in services
6. Market Access in Goods: Provides members no market access in goods/Provides members full market access in goods
7. Internal tariff reduction: Member countries have substantially increased tariffs within the PTA/Member countries have substantially reduced tariffs within the PTA
8. Intra-PTA trade: The PTA has not contributed to increased trade/The PTA has contributed to increased trade among members
9. Dispute Settlement: No formal channel for dispute settlement/PTA provides a formal channel for dispute settlement that is easy for members to use
10. Monitoring: Weakly monitors actions of its members, with no formal channels for monitoring/Strongly monitors actions of its members, through formal channels
11. Enforcement: Does not enforce its own rules and resolutions; does not punish non-compliance/Strongly enforces its own rules and resolutions; consistently punishes noncompliance
12. Legalization: Low levels of legalization and formalization in founding treaties and subsequent documents/High levels of legalization and formalization in founding treaties and subsequent documents
13. Escape Clauses: Difficult for members to invoke escape clauses/Easy for members to invoke escape clauses
14. Political Influence: No political influence on members and region/High degree of political influence on members and region

15. Ambition of Goals: The PTA's goals are limited/The PTA sets high, ambitious goals
16. Scope of Goals: The agenda items over which PTA members negotiate do not go beyond traditional trade issues/The agenda items over which PTA members negotiate go beyond traditional trade issues
17. Ability to Meet Own Goals: Does not meet own goals and targets in terms of market access/ Completes agreed-on goals and targets, in terms of market access, on schedule and to the fullest extent possible
18. Delegation: The PTA grants no authority to third parties to implement, interpret and apply rules/The PTA grants high levels of authority to third parties to implement, interpret and apply rules
19. Obligation: The agents of the PTA are not bound by legal commitment/The agents of the PTA are bound by a high level of legal commitment
20. Precision: The rules set by the PTA are unambiguous, leaving little scope for interpretation/The rules set by the PTA are highly ambiguous, leaving much scope for interpretation
21. Rules of Origin: The agreement does not include provisions for rules of origin/The agreement includes provisions for rules of origin
22. WTO Compliance: The PTA does not follow restrictions and regulations for PTA WTO compliance/The PTA follows restrictions and regulations for PTA WTO compliance
23. Nontariff Barriers: No nontariff barriers among members of the PTA/Nontariff barriers proliferate among members of the PTA
24. Strong-Weak: On a general scale of weak to strong, taking all aspects of PTA policy into account, how would you place this PTA? Weak/Strong

Table 1: Expert Survey: Descriptive Statistics for RTAs

RTA	Acronym	Number of				
		Responses	Mean	SD	Minimum	Maximum
Andean Common Market	AC	150	5.40	2.15	1	10
Latin American Integration Association	ALADI	70	7.60	2.80	1	10
Arab Maghreb Union	AMU	38	3.75	2.99	1	10
Asia-Pacific Economic Cooperation	APEC	73	4.11	2.60	1	9
Association of Southeast Asian Nations	ASEAN	258	5.23	2.34	1	10
Bangkok Agreement	BA	70	1.84	1.74	1	9
Central American Common Market	CACM	56	4.90	1.62	4	7
Caribbean Community	CARICOM	224	6.41	2.18	1	10
Central European Free Trade Agreement	CEFTA	98	7.05	1.97	2	10
Economic and Monetary Community of Central Africa	CEMAC	48	5.36	2.81	1	10
Commonwealth of Independent States	CIS	60	4.65	2.97	1	10
Common Market for Eastern and Southern Africa	COMESA	47	7.81	2.04	2	10
China-ASEAN FTA	China-Asean	10	3.70	1.89	2	7
East African Community	EAC	25	7.44	2.36	1	10
Economic Community of Central African States	ECCAS	49	5.52	2.93	1	10
Economic Community of West African States	ECOWAS	82	6.68	2.84	1	10
European Economic Area	EEA	21	8.10	2.47	1	10
European Free Trade Agreement	EFTA	278	7.35	2.28	1	10
Economic Partnership Agreement	EPA	41	7.46	1.61	1	9
European Union	EU	492	8.97	2.10	1	10
Eurasian Economic Community	EURASEC	21	5.00	2.49	1	8
Greater Arab Free Trade Area	GAFTA	36	5.61	2.89	1	10
Gulf Cooperation Council	GCC	169	6.17	2.43	1	10
India-Singapore FTA	India-Singapore	2	2.00	0.00	2	2
Japan-Mexico FTA	Japan-Mexico FTA	3	5.00	4.58	1	10
Korea-Chile FTA	Korea-Chile FTA	7	7.14	1.68	5	10
Korea-US FTA	KORUS FTA	27	5.59	2.48	2	10
Mano River Union	MRU	8	1.50	1.41	1	5
Melanesian Spearhead Group	MSG	62	5.50	2.54	1	10
Common Market for the South	Mercosur	326	6.28	2.39	1	10
North American Free Trade Agreement	NAFTA	265	7.25	2.12	1	10
Organization of East Caribbean States	OECS	83	7.15	2.03	1	10
Pacific Island Countries Trade Agreement	PICTA	34	6.21	1.98	1	9
South Asian Association for Regional Cooperation	SAARC/SAFTA	63	4.37	2.16	1	8
South African Customs Union	SACU	100	6.29	2.59	1	10
South African Development Community	SADC	138	6.45	2.36	1	10
Shanghai Cooperation Organization	SCO	21	3.81	2.62	1	9
South Pacific Trade and Economic Co-Operation Agreement	SPARTECA	48	2.19	1.93	1	9
Singapore-Japan FTA	Singapore-Japan	2	2.00	0.00	2	2
The West African Economic and Monetary Union	UEMOA	74	6.60	2.81	1	10
<b>Total</b>		<b>3831</b>	<b>6.47</b>	<b>2.80</b>		



Table 2: **Expert Survey: Descriptive Statistics for RTA Dimensions (Means)**

Variable	N	Mean	Std. Dev.	Min	Max
Ability	33	4.71	2.48	1	9.5
Ambition	31	6.89	2.28	1	10
Delegation	28	5.12	2.84	1	10
DSM	28	4.18	2.72	1	9
Enforcement	27	3.72	2.40	1	9.5
Escape Clauses	27	6.25	2.35	1	10
Internal Tariff Reduction	33	6.27	2.46	1	10
International Influence	30	3.64	2.00	1	9.8
IntraPTA Trade	32	6.06	2.25	2	10
Legalization	28	6.17	2.48	1	10
Goods	35	5.77	2.31	1	10
Services	33	3.47	2.58	1	9.5
Monitoring	28	4.33	2.46	1	9.6
NTBs	32	5.59	2.47	1	10
Obligation	21	6.54	2.48	1	9.75
Political Influence	31	5.48	2.37	1	9.6
Precision	29	5.60	2.29	1	10
Secretariat	31	4.68	2.44	1	9.6
Rules of Origin	12	4.90	3.49	1	10
Scope	33	6.29	2.55	1	10
Tariff Reduction	4	6.75	2.99	3	10
Trade Diversion	27	4.36	2.07	1	8
Negotiator Capacity	30	5.33	2.34	1	9.2
WTO Compliance	12	7.13	2.26	3	9

Table 3: **Expert Survey: Descriptive Statistics for RTA Dimensions (Salience)**

Variable	N	Mean	Std. Dev.	Min	Max
Ability	33	6.78	1.78	3	10
Ambition	31	7.18	1.87	3	10
Delegation	25	5.63	2.61	1	10
DSM	30	6.03	2.73	1	9.2
Enforcement	27	4.96	2.49	1	9.25
Escape Clauses	24	6.40	2.69	1	9
Internal Tariff Reduction	31	7.55	2.26	1	10
International Influence	29	5.18	2.43	1	9.75
IntraPTA Trade	33	7.13	2.19	2	10
Legalization	28	6.65	2.42	1	10
Goods	35	7.09	2.32	1	10
Services	32	5.28	2.66	1	9.25
Monitoring	28	4.94	2.64	1	9.4
NTBs	32	5.59	2.45	1	9.3
Obligation	20	7.30	2.21	1	9.5
Political Influence	31	6.59	1.99	2	9.5
Precision	29	5.78	2.46	1	9.5
Secretariat	28	6.82	2.13	1	10
Rules of Origin	12	4.90	3.66	1	10
Scope	32	6.96	2.28	2	10
Tariff Reduction	4	8.50	1.29	7	10
Trade Diversion	28	2.79	1.75	1	7
Negotiator Capacity	30	6.35	2.30	1	9.6
WTO Compliance	12	6.25	2.49	1	10

Table 4: Correlation Matrix for Dimensions

	Ability	Ambition	Delegation	DSM	Enforcement	Escape	Tariff	Influence	IntraPTA	Legalization	Goods	Services	Monitoring	NTBs	Obligation	Political	Precision	Secretariat
Ability	1.00																	
Ambition	0.31	1.00																
Delegation	0.40	0.99	1.00															
DSM	1.00	0.38	-0.24	1.00														
Enforcement	0.79	-0.34	-0.97	0.81	1.00													
Escape	-0.61	-0.94	-0.97	-0.59	0.00	1.00												
Low Tariff	0.89	-0.16	-0.06	0.90	0.98	-0.18	1.00											
Intl Influence	-0.32	0.80	0.74	-0.34	-0.83	-0.56	-0.72	1.00										
IntraPTA Trade	0.53	-0.64	-0.56	0.74	0.94	0.35	0.86	-0.97	1.00									
Legalization	0.99	0.18	0.28	0.99	0.87	-0.50	0.94	-0.44	-0.44	1.00								
Goods	0.95	-0.02	0.09	0.95	0.95	-0.32	0.99	-0.61	-0.61	0.78	1.00							
Services	1.00	0.32	0.42	1.00	1.00	0.78	-0.62	0.88	-0.30	0.52	0.99	1.00						
Monitoring	0.16	0.99	0.97	0.14	0.14	-0.47	-0.88	-0.31	0.88	-0.75	0.03	-0.17	1.00					
NTBs	0.42	-0.74	-0.67	0.44	0.44	0.88	0.47	0.79	-0.99	0.99	0.54	0.69	0.40	1.00				
Obligation	0.84	-0.25	-0.15	0.86	0.86	1.00	-0.09	1.00	-0.78	0.90	0.91	0.97	0.83	-0.39	1.00			
Political	-0.35	0.78	0.71	-0.38	-0.38	-0.85	-0.53	-0.74	1.00	-0.98	-0.48	-0.64	-0.34	0.87	-1.00	1.00		
Precision	0.05	-0.94	-0.89	0.07	0.07	0.65	0.76	0.50	-0.96	0.87	0.18	0.37	0.03	-0.98	0.93	0.58	1.00	
Secretariat	1.00	0.26	0.35	1.00	1.00	0.82	-0.57	0.91	-0.37	0.58	1.00	0.96	1.00	0.11	0.46	0.87	-0.95	1.00
ROO	0.29	-0.82	-0.76	0.31	0.31	0.82	0.58	0.70	-1.00	0.97	0.42	0.59	0.28	-0.90	0.99	0.76	-1.00	0.10
Scope	0.64	0.93	0.96	0.63	0.63	0.04	-1.00	0.22	0.52	-0.30	0.53	0.36	0.66	0.86	-0.43	0.13	0.49	-0.73
Tariffs	0.76	-0.38	-0.28	0.78	0.78	1.00	0.05	0.97	-0.85	0.95	0.84	0.93	0.75	-0.51	0.90	0.99	-0.87	0.68
Diversification	-0.84	0.26	0.16	-0.85	-0.85	-1.00	0.08	-1.00	0.78	-0.91	-0.91	-0.97	-0.83	0.40	-0.84	-1.00	0.81	-0.58
Negotiator	0.87	0.74	0.81	0.85	0.85	0.38	-0.93	0.54	0.20	0.04	0.79	0.66	0.87	0.64	-0.10	0.46	0.16	-0.46
WTO	0.94	-0.02	0.08	0.95	0.95	0.95	-0.31	0.99	-0.61	0.78	0.98	1.00	0.94	-0.17	0.69	0.97	-0.64	0.38

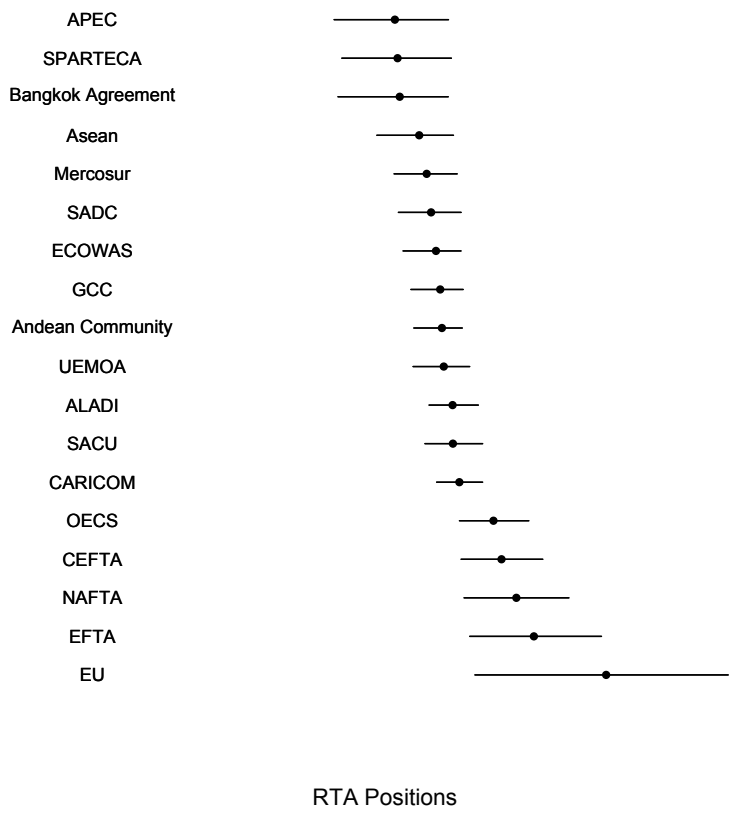


Figure 1: RTA Factor Scores,  $\phi$

Table 5: **Factor Analysis Scores for Dimensions**

	Mean	SD	Naive SE	Time-series SE
$\lambda$ Ability	-3.37	3.90	0.12	0.37
$\lambda$ Ambition	-2.27	2.93	0.09	0.27
$\lambda$ Delegation	-2.39	2.85	0.09	0.26
$\lambda$ DSM	-2.22	2.72	0.09	0.24
$\lambda$ Enforcement	-2.92	3.82	0.12	0.34
$\lambda$ Goods	-3.17	3.60	0.11	0.34
$\lambda$ IntraPTA Trade	-3.32	3.93	0.12	0.38
$\lambda$ Legalization	-3.14	3.65	0.12	0.34
$\lambda$ Monitoring	-3.16	3.73	0.12	0.32
$\lambda$ NTBs	-0.82	1.29	0.04	0.06
$\lambda$ Scope	-2.18	2.53	0.08	0.23
$\lambda$ Secretariat	-3.21	4.13	0.13	0.37
$\lambda$ Services	-3.00	3.22	0.10	0.29
$\lambda$ Internal Tariff Reduction	-3.10	3.59	0.11	0.31
$\lambda$ Trade Diversion	-0.34	1.66	0.05	0.05
$\psi$ Ability	0.19	0.09	0.00	0.00
$\psi$ Ambition	0.67	0.27	0.01	0.01
$\psi$ Delegation	0.63	0.23	0.01	0.01
$\psi$ DSM	0.68	0.28	0.01	0.01
$\psi$ Enforcement	0.41	0.17	0.01	0.01
$\psi$ Goods	0.28	0.12	0.00	0.00
$\psi$ IntraPTA Trade	0.23	0.11	0.00	0.00
$\psi$ Legalization	0.30	0.13	0.00	0.00
$\psi$ Monitoring	0.28	0.12	0.00	0.00
$\psi$ NTBs	1.01	0.37	0.01	0.01
$\psi$ Scope	0.68	0.26	0.01	0.01
$\psi$ Secretariat	0.28	0.11	0.00	0.00
$\psi$ Services	0.36	0.15	0.00	0.00
$\psi$ Internal Tariff Reduction	0.33	0.15	0.00	0.00
$\psi$ Trade Diversion	1.07	0.40	0.01	0.01

Table 6: **Factor Analysis Scores for RTAs**

	Mean	SD	Naive SE	Time-series SE
$\phi$ ALADI	0.04	0.08	0.00	0.00
$\phi$ APEC	0.43	0.21	0.01	0.01
$\phi$ Andean Community	0.11	0.09	0.00	0.00
$\phi$ Asean	0.27	0.14	0.00	0.01
$\phi$ Bangkok Agreement	0.40	0.20	0.01	0.01
$\phi$ CARICOM	-0.01	0.08	0.00	0.00
$\phi$ CEFTA	-0.30	0.15	0.00	0.01
$\phi$ ECOWAS	0.15	0.11	0.00	0.00
$\phi$ EFTA	-0.52	0.24	0.01	0.01
$\phi$ EU	-1.01	0.45	0.01	0.02
$\phi$ GCC	0.12	0.10	0.00	0.00
$\phi$ Mercosur	0.21	0.12	0.00	0.01
$\phi$ NAFTA	-0.40	0.19	0.01	0.01
$\phi$ OECS	-0.24	0.13	0.00	0.01
$\phi$ SACU	0.04	0.10	0.00	0.00
$\phi$ SADC	0.18	0.12	0.00	0.01
$\phi$ SPARTECA	0.41	0.20	0.01	0.01
$\phi$ UEMOA	0.10	0.10	0.00	0.00