“Does Institutional Design Matter?
A Study of Trade Effectiveness and PTA Flexibility/Rigidity”

by

David H. Bearce
Professor of Political Science/International Affairs
University of Colorado, Boulder
David.Bearce@colorado.edu

Cody D. Eldredge
Ph.D. Candidate, Department of Political Science
University of Colorado, Boulder
Cody.Eldredge@colorado.edu

Brandy J. Jolliff
Ph.D. Candidate, Department of Political Science
University of Colorado, Boulder
Brandy.Jolliff@colorado.edu

Prepared for IPES 2014

Abstract: This paper examines the trade effect of flexibility design features within preferential trading arrangements (PTAs). Using a gravity model of bilateral trade that incorporates multilateral trade resistance, it reports three main results. First, it finds that unconstrained escape provisions are bad for PTA trade effectiveness. Second, adding some restrictions to these escape provisions more than offsets the negative effect of unconstrained escape, leading to PTAs that are more trade effective than those without escape options. Third, adding more restrictions beyond a certain point only makes PTAs less trade effective. Thus, while too much institutional flexibility is bad for trade, so is too much institutional rigidity.

Acknowledgements: Earlier drafts were presented at ISA 2014 in Toronto, at the Institute of Behavioral Sciences at CU Boulder, and at APSA 2104 in Washington, DC. Our biggest thanks go to Jeff Kucik for providing his data on PTA design. We also thank Lee Alston, Leo Baccini, Manfred Elsig, Julia Gray, Moonhawk Kim, Joe Jupille, Jeff Kucik, Keith Maskus, Sara Mitchell, and Jamie Scalera for their helpful comments. Finally, we are grateful to Karen Naranjo for her research assistance.

The research program on the rational design of international institutions (RDII) represents a useful starting point for the arguments and evidence presented in this paper. In the RDII research program, scholars have explored what factors determine the variation in international institutional design with a particular focus on explaining the different features related to membership, scope, centralization, control, and flexibility (Koremenos, Lipson, and Snidal 2001).

In treating these design features as dependent variables, RDII scholars have effectively assumed their rationality. Stated differently, the choice for or against a particular design feature is assumed to be rational in the sense that it has some causal effect on institutional performance. Thus, policymakers who want to achieve a particular outcome choose the design features that facilitate this outcome. An alternative explanation is that these design features are not chosen so deliberately. Instead, policymakers creating new institutions may simply select design features that already exist in similar institutional types even when these design features have no meaningful effect on institutional performance. Indeed, it is not clear how policymakers could know with any great confidence whether these design features have any causal effect on institutional performance.

This rationality assumption stands as a primary, if not *the* primary, criticism of the RDII research program. As Wendt (2001, 1043) queried on this point, "what if [such] design features are not, in fact, functional [i.e., effective in achieving the desired ends]?" Following the challenge issued by Martin and Simmons (1998) that scholars can only move forward by focusing on *how* institutions matter (i.e., by treating institutions as independent variables), Duffield (2003, 414) advanced a similar critique: "[RDII] moves in the opposite direction, revisiting international institutions as dependent variables." On this basis, one could argue that it makes little sense to study the variation in institutional design until it has been demonstrated that these design features have an effect on important international outcomes. Without such a demonstration, one risks elevating meaningless treaty details and exaggerating the importance of irrelevant institutional features.

The paper thus seeks to address the prior question of institutional design effectiveness with a particular focus on flexibility features within preferential trading arrangements (PTAs): does PTA flexibility help or hinder international trade? We justify this focus based on the understanding that flexibility has been a central concept in the RDII research program (see the various papers in Kormenos, Lipson, and Snidal 2001; Kormenos 2001; Kormenos 2005) with PTAs as a proliferating institutional form in the international trade system given the difficulties associated with deepening the GATT/WTO (Mansfield 1998; Mansfield and Reinhardt 2003). Furthermore, scholars have advanced arguments about the effect of trade flexibility in both directions. Flexibility features allow member-states to escape their commitments, potentially hindering the growth of international trade. Alternatively, flexibility features may reassure governments, facilitating the growth of international trade by allowing them to enter into deeper agreements.

Despite the theoretical interest in trade flexibility, there exists little evidence bearing on these competing arguments about its effect. Identifying a domestic antidumping mechanism as a flexibility indicator, Kucik and Reinhardt (2008) find that countries with an antidumping mechanism are more likely to join the General Agreement on Tariffs and Trade (GATT) and the successor World Trade Organization (WTO), have lower binding tariffs when they join, and maintain lower applied tariffs within this global trade regime. However, identifying binding tariff overhang (binding minus applied tariffs) within the GATT/WTO as a flexibility feature, Pelc (2013) shows that more overhang has a strong negative effect on international trade.

Given these different (even contrary) results about the effect of trade flexibility within the GATT/WTO, it becomes important to examine the effect of flexibility within PTAs, the alternative institutional form towards freer trade. Indeed, there is arguably more flexibility variation among these regional arrangements than within the global trade regime where most of the flexibility features apply to all of the member-states. Focusing on these regional arrangements, Kucik (2012) demonstrated the variation in PTA flexibility/rigidity and then explored the determinants of the same using an operational measure that includes PTA-specific restrictions on the use of safeguards, antidumping, and countervailing duties.

We make use of Kucik's rigidity measure as an independent variable, coding the variety of PTA flexibility/rigidity into a gravity model of international trade, while controlling for other PTA design features including depth, width, and legalism. In a series of models with increasing restricted samples, we report three main results. First, unconstrained escape provisions are bad for trade effectiveness. But second, adding some restrictions to these escape provisions more than offsets the negative effect of unconstrained escape, thus leading to PTAs that are more trade effective than those without escape options. Third, adding more restrictions beyond a certain point only serves to make PTAs less trade effective. In this regard, while too much institutional flexibility is bad for trade, so is too much institutional rigidity.

1. Arguments about Trade Agreement Flexibility

Before presenting the arguments for and against trade agreement flexibility, it is important to first consider two different conceptions of "flexibility" in the context of international agreements and then briefly review the existing evidence on PTA trade effectiveness. This first preliminary exercise will identify different operational indicators of PTA flexibility/rigidity. The second will provide some context for understanding the potential effects associated with PTA variation in this dimension.

Conceptions of Agreement Flexibility

To define agreement flexibility with an application to trade agreements, it is important first to consider two different conceptions of flexibility as it relates to international agreements, more broadly. The first conception assumes that international agreements are fairly comprehensive and generally enforceable once they are negotiated. This perspective, with its origins in international law (e.g. Chayes and Chayes 1993), observes that international agreements are often quite long and detailed (although not necessarily complete contracts). Furthermore, it expects that signatory governments largely intend to comply with the agreement; otherwise, they would not sign it. From this understanding, agreement flexibility comes in the form of opt-out provisions, which allow signatory governments to escape temporarily their treaty commitments while remaining inside the institution.

The classic example of an opt-out, or escape, clause within a detailed international trade agreement is Article XIX: Emergency Action on Imports of Particular Products in the General Agreement on Tariffs and Trade (1947). This so-called "safeguard" provision allowed member-states "to suspend the obligation in whole or in part" in response to "any product [that] is being imported into the territory of that contracting party in such increased quantities and under such conditions as to cause or threaten serious injury to domestic producers". Some version of this basic safeguard provision has been inserted into a large majority of the PTAs that have been notified to the GATT/WTO under Article XXIV or the Enabling Clause. Likewise, many PTAs now include antidumping and countervailing duty provisions, which also allow member-governments to temporarily raise trade barriers, while remaining inside (and compliant with) the agreement. From this first conception, trade agreements with more such escape provisions are more flexible, while those with less are more rigid.

The second conception assumes that international agreements are limited in coverage and difficult to enforce. Even when they are long and detailed, international agreements remain as incomplete contracts. And even if signatory governments intend to comply with their agreements when they sign them, they will always face future unexpected shocks that create an incentive for non-compliance. From this understanding, international agreements, including trade agreements, are inherently flexible. Agreement rigidity, to the extent that it can be created at all, comes from explicitly limiting what governments can do, or in the written restrictions on how governments can use their various escape options. Indeed, if most PTAs include escape provisions, then arguably most of the flexibility/rigidity variation comes from these written restrictions.

As mentioned in the introduction, Kucik (2012) coded the variation in PTA flexibility/rigidity based on the written restrictions related to safeguard, antidumping and countervailing duty provisions (e.g. requiring the prior demonstration of injury, limits to their duration, and limits to the amount of trade restriction that can be imposed). Following this second conception, trade agreements with fewer such restrictions on the use of their escape provisions can be identified as more flexible, while those with greater restrictions are more rigid.

From these two different conceptions, trade agreement flexibility/rigidity can be identified by two related indicators: 1) flexibility based on the presence (and number) of escape provisions and 2) rigidity based on the amount of restrictions associated with the use of such escape provisions. It should be noted that the second indicator is nested within the first as there can be no formal restrictions on the use of an escape provision that does not exist.

Having identified the operation indicators for PTA flexibility/rigidity, it is also important to note that there is much empirical evidence showing that PTAs have a strong positive effect on bilateral trade. Indeed, when using the gravity model of trade (to be described below), the estimated impact of PTAs on dyadic trade is often larger than that of the GATT/WTO. The surprisingly weak effect of the GATT/WTO has been a subject of intense debate (e.g. Rose 2004; Tomz, Goldstein and Rivers 2007), but these studies and others (e.g. Baier and Bergstrand 2007) consistently show a strong average PTA effect when using a gravity model specification. However, this strong result certainly does not mean that all bilateral/regional arrangements are equally effective at increasing trade. Studies that estimate the effect of different PTAs (e.g. Kono 2007) demonstrate that even if the average effect is large and positive, there is much variation in the trade impact of these bilateral/ regional trade institutions. Perhaps PTA flexibility/rigidity can help explain this variation.

Hypotheses about Trade Agreement Flexibility

What might be termed the "conventional wisdom" (Dunoff 2010, 402) about trade agreement flexibility/rigidity proposes that greater flexibility is bad for trade effectiveness. If the real purpose of a PTA is to reduce existing trade barriers as a means to increasing cross-border exchange, then escape provisions and opt-out clauses that allow member-governments to raise trade restrictions run directly contrary to this goal. As Helfer (2013, 186) summarized on this point: "After a treaty enters into force…, escape clauses may have deleterious consequences for international cooperation. The most basic concern is that escape mechanisms authorize deviant behavior precisely when treaty compliance is needed most."

This conventional wisdom offers a linear hypothesis about the relationship between agreement rigidity and its trade effectiveness, as illustrated in Figure 1 with trade effectiveness on the y-axis and agreement flexibility/rigidity on the x-axis. At the far left of the x-axis are the most flexible agreements, or those with generous escape provisions that have no restrictions on their use. The conventional wisdom expects that such highly flexible PTAs should have a strong negative impact on international trade (compared to agreements that lack such escape provisions). While it may be possible to reduce this negative effect by making the escape provisions more restrictive (thus moving right on the x-axis towards greater rigidity), such restrictions can only mitigate the negative effect associated with escape provisions. Thus the upward sloping line for the conventional wisdom does not cross over into positive (+) trade effectiveness.

[Figure 1 here]

An alternative proposition has been termed the "flexibility hypothesis" (Kucik and Reinhardt 2008, 478). Central to this hypothesis is the understanding that governments face a time-inconsistency problem with regards to international trade: in the present, they want to liberalize their domestic market to realize the gains from trade, but in the future, they may also face an economic shock that makes returning to trade restrictions a more politically attractive option. Escape provisions allow governments to respond to this shock with temporary trade restrictions and then to return to full compliance within the free trade agreement at some later date. Facing this time-inconsistency problem, flexibility may be an attractive feature, inducing governments to sign deeper agreements than they would otherwise (Johns 2014).

However, it is important to understand that the flexibility hypothesis does not propose that more flexibility always results in greater trade effectiveness, or the inverse of the conventional wisdom. Like the conventional wisdom, the flexibility hypothesis expects escape provisions, when they are exercised, to have a negative, or non-cooperative, effect on international trade (Rosendorff 2005). But if escape options allow governments to sign deeper trade agreements, then some of the negative effect associated with the defection of exercised escape may be offset by deeper cooperation within the agreement when escape is not exercised. Hence, the flexibility hypothesis is more sanguine about negative effect of unrestricted escape (the far left side of the x-axis) than is the conventional wisdom, as illustrated in Figure 1.

To the extent that there is some cost to, or restrictions on, the use of these escape provisions, then non-cooperation should be reduced with temporary escape becoming a rarer event, thus resulting in greater trade effectiveness through increased cooperation within the agreement. Thus for at least some range of the flexibility/rigidity dimension, the flexibility hypothesis (like the conventional wisdom) expects greater trade effectiveness with more escape restrictions. Indeed, this positive relationship should be even stronger than expected by the conventional wisdom since greater cooperation (or reduced defection) also occurs within a deeper trade agreement, a development that is not anticipated by the conventional wisdom.

But if these escape restrictions are seen as too numerous or onerous, then governments may become unwilling to sign deeper agreements, thus resulting in reduced trade effectiveness through shallower agreements even as escape defections become rarer events. As Rosendorff and Milner (2001) wrote about costly escape provisions within free trade agreements: the cost for temporary escape "must not be too high or it will eliminate any flexibility and make the [trading] system unstable; but it must also not be too low or it will render 'cooperation' ineffective" (ibid 835). Helfer (2013, 176) likewise noted about international agreements more broadly: "a principal challenge facing treaty negotiators is to select a suite of appropriately constrained flexibility mechanisms that facilitate agreement among states *ex ante* while deterring opportunistic uses of those mechanisms *ex post* after the treaty enters into force." Per the flexibility hypothesis, too much flexibility can be expected to hurt agreement effectiveness (consistent with the conventional wisdom), but so can too much rigidity (unlike the conventional wisdom).

Before proceeding to the empirical tests, it is worthwhile to consider the null hypothesis as something more than just a statistical default. The obvious null hypothesis is that these flexibility/ rigidity indicators have no real effect on trade agreement effectiveness. Indeed, World Polity theory (e.g. Meyer et al. 1997) would argue that PTAs and their design features are chosen primarily for symbolic reasons and not because they have any causal impact on trade.[1] Consistent with this expectation, Jo and Namgung (2012) provide evidence of design feature emulation among PTAs, with Gray and Slapin (2012) offering evidence to suggest that formal design features may not matter much for actual PTA effectiveness.

Specifically with reference to flexibility/rigidity features considered here, the logic behind the conventional wisdom is based on greater defections with flexibility in the form of escape provisions. But if international agreements are inherently flexible (consistent with the second conception discussed above), then even PTAs without escape clauses should experience frequent defections. On this basis, the presence or absence of escape provisions should not matter much for PTA effectiveness. Likewise, the logic behind the flexibility hypothesis assumes that escape options with moderate restrictions

[1] For an application of this logic to PTAs, see Jupille, Jolliff, and Wojcik (2013).

should induce deeper trade agreements. But empirical evidence on this underlying proposition remains somewhat limited.[2]

2. Evidence about Trade Effectiveness and PTA Flexibility/Rigidity

We use a gravity model of international trade that incorporates multilateral trade resistance to test these hypotheses about the different possible relationships between trade effectiveness and PTA flexibility/rigidity.[3] With a gravity model specification, our operational measure for trade effectiveness becomes the amount of trade to country *i* from country *j* in year *t* (e.g., *i*'s *imports* from *j* in *t*). We recognize that there are other possible ways to measure trade effectiveness, such as the amount of trade restrictions (of various forms including, but not limited to, tariffs), but reduced barriers to trade are only an intermediate step towards increasing actual trade.

The Gravity Model Specification

To measure the amount of trade to country *i* from country *j* in year *t*, we begin with the bilateral import data available from the International Monetary Fund's (IMF) *Direction of Trade Statistics*. Where the import data are missing, we use the corresponding export measure (from country *j* to country *i*) multiplied by 1.1 (to take account of the cost of freight and insurance).[4] These IMF data come in current US dollars, which we converted to constant 2009 US dollars using the deflator from the US

[2] For some evidence supportive of this proposition, see Baccini, Dur, and Elsig 2013.

[3] Introduced by Tinbergen (1962), the gravity framework has become the standard model for analyzing the determinants of bilateral trade. Its basic logic is that bilateral trade increases with income and decreases with distance and other barriers to exchange. Some theoretical foundations for the gravity model are provided by Anderson (1979) and Anderson and van Wincoop (2003).

[4] When there remain missing values within the dyadic time-series, we interpolate the missing *internal* values, but we do not extrapolate any values either before or after the first and last data point from the IMF data for each dyadic time-series.

Department of Commerce, Bureau of Economic Analysis. Given our gravity model specification, we take the logged value of *Imports* as our dependent variable, as shown in equation (1).

$$\ln(\mathit{Imports})_{ijt} = B_0 + B_1\ln(\mathit{ProductGDP})_{ijt} + B_2\ln(\mathit{Distance})_{ij} + B_3\mathit{GATT/WTO}_{ijt}$$
$$+ B_4\mathit{PTA}_{ijt} + B_5\mathit{PTA\ Design}_{ijt} + \mathit{Dyad}_{ij} + \mathit{Year}_t + \mathit{Importer\text{-}Year}_{it} + \mathit{Exporter\text{-}Year}_{jt} + \varepsilon_{ijt} \quad (1)$$

On the right-hand of the equation, the standard gravity model specification begins with the logged value of the product of GDP for the two countries in the directed dyad in year *t*. The standard gravity specification also includes a measure of *Distance* between the two countries and sometimes other dyad-specific features such as whether (or not) they share a common language, had a colonial relationship, have a land border, and include landlocked or island countries. But given our use of directed dyad fixed effects (*Dyad*) to deal with selection issues (e.g., countries are not randomly assigned into PTAs with different design features), *Distance* and any other dyad-specific variables drop from our estimates. Given the number of possible factors that might lead to selection bias, we privilege the use of dyad fixed effects over these directly measurable, but limited, dyadic variables, which are of no particular theoretic interest at least in this paper.

In using *Importer-Year* and *Exporter-Year* fixed effects to incorporate the concept of multilateral trade resistance,[5] ln(*ProductGDP*) will also drop from our estimates due to collinearity with these two additional sets of fixed effects. Once again, we privilege the use of fixed effects to capture multilateral trade resistance since we have no theoretical interest in the effect of income/market size on trade. The gravity model specification often includes *Year* fixed effects, but this fourth set of fixed effects will drop from our estimates due to perfect collinearity with the *Importer-Year* and *Exporter-Year* fixed effects. It

[5] The concept of "multilateral trade resistance" comes from the understanding that the amount of trade between two countries is not simply based on the barriers to exchange between them (such as distance), but also on the average barriers to trade with other potential trading partners (Anderson and van Wincoop 2003). Since these average barriers to trade are hard to measure directly, but are specific to the trading country with variation over time, Baldwin and Taglioni (2006) showed how they can be approximated with importer-year and exporter-year fixed effects.

should be noted that these fixed effects are conditioned out of the data, rather than estimated directly.[6] As such, and given their number (over 30,000 for $Dyad_{ij}$ and approximately 10,000 for both $Importer\text{-}Year_{it}$ and $Exporter\text{-}Year_{jt}$ when using our full sample), these three sets of fixed effects cannot be reported in our statistical tables.

Given our general interest in the trade effectiveness of international trade institutions, our gravity model specification also includes a dichotomous variable for joint membership *and participation* in the *GATT/WTO*, following Tomz, Goldstein and Rivers (2007). And while we are most interested in the effect of PTA design features, our specification includes a dichotomous variable for joint *PTA* membership, which will pick up on *other* PTA effects once the specific PTA design variables get inserted into our gravity model. Based on our interest in the effects of PTA flexibility/rigidity and our use of Kucik's (2012) measure for this concept (to be described below), our PTA sample consists of those coded in his dataset, which are taken from the World Trade Organization's Regional Trade Agreements Information System and includes both "in force" agreements and now "inactive" agreements (that were once in force).[7] Kucik's dataset included over 300 PTAs, which represents a smaller set of PTAs than found in Dur, Baccini, and Elsig (forthcoming), but a larger set than used in Baier and Bergstrand (2007).

It should be noted that to reduce the problems associated with overlapping PTA membership, which could be serious given that many of PTAs in our sample are "association" agreements involving the European Union (EU) and the European Free Trade Association (EFTA), we treat both of these PTAs as effective countries, summing their trade for every year of their existence. This decision helps ensure that the design features of the EU and EFTA are treated separately from those of their association agreements. This decision is also consistent with the WTO's treatment of these association

[6] For this fixed effects estimation, we use the *reghdfe* command in Stata.

[7] http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx .

agreements as "Bilateral" where "One Party is an RTA". Although the Association of Southeast Asian Nations (ASEAN) and the Caribbean Community (Caricom) have fewer such agreements, we do the same for their association agreements.

Before inserting any PTA design variables, we estimate the ln(*Imports*) model just described using all directed dyad/years for the period 1950-2010 in order to assess the average *PTA* effect. This baseline estimate is presented as model 1 in Table 1. Without any of these design variables, *PTA* is statistically significant, taking on a positive sign of 0.43. Given our log linear specification, this coefficient indicates that joint membership in a preferential trading arrangement has the effect of increasing trade by about 54 percent compared to a dyad where the countries lack this joint membership ($e^{0.43} - 1 = 0.54$).[8] This estimate accords closely with the median value (0.47) reported by Head and Mayer (2014, 160) in their meta-analysis of PTA estimates using the gravity model specification. So we judge that our PTA coverage provides a reasonable baseline estimate for the PTA effect on bilateral trade, and we thus proceed using this set of bilateral/regional trading arrangements.

PTA Design Variables

Into this model, we now add a series of PTA design variables. To the best of our knowledge, this is the most fully specified model in terms of PTA design that has yet appeared in the literature. These design variables are all coded as 0 for the dyad/years without PTA coverage since they lack a PTA for which these design features might be present.[9] To avoid sample selection bias, our model will

[8] This PTA effect can also be compared to that of the *GATT/WTO*, which takes on a statistically significant coefficient of 0.21, indicating that joint membership/participation in the global regime has increased trade by about 23 percent ($e^{0.21} - 1 = 0.23$). Our result is smaller than the effect reported in Goldstein, Rivers, and Tomz (2007) or in Tomz, Goldstein, and Rivers (2007). But it is larger than the GATT/WTO effect reported by Dur, Baccini, and Elsig (forthcoming).

[9] Since these PTA design variables are internal to, or nested within, the *PTA* variable (i.e., coded as non-zero only when *PTA*=1), they cannot be entered as multiplicative terms (i.e., interacted with *PTA*). If entered as multiplicative terms, then either the *PTA Design* constitutive term or the *PTA**PTA Design*

first be estimated using a statistical sample that includes all PTA and non-PTA directed dyad/year observations for which we have data; this full sample also allows for a comparison among the PTA observations with the design features, the PTA observations that lack the design features, and the non-PTA observations that also lack the design features. But our statistical model will also be estimated using reduced samples that 1) eliminate the dyad/year observations associated with small values of trade, 2) include only PTA dyad/year observations, and 3) use the PTA as the unit of analysis.

The PTA design variables begin with controls for both depth and width. For the former, we use the 8-point measure of *PTA Depth* presented in Dur, Baccini, and Elsig (forthcoming).[10] For the latter, we include a dummy variable indicating a *Plurilateral PTA* based on the WTO's "RTA Composition" coding: plurilateral or bilateral (including bilateral; one party is an RTA).[11] Given the expectation that both deeper and wider PTAs should be more trade effective (although recognizing a potential tradeoff between depth and width), we anticipate a positive sign both of these controls.

Next we include a dummy variable indicating whether or not the PTA includes dispute settlement mechanism of any type.[12] Following Kono's (2007) results, we expect *PTA DSM* to take on a positive coefficient, indicating that more legalistic PTAs are associated with greater bilateral trade.

To capture the various possibilities associated with PTA flexibility/rigidity discussed in the previous section, we include three independent variables. The first is *PTA Escape*, which counts the

interaction term must drop from the model since the two are perfectly collinear. This can be illustrated with the three observation dataset below.

	PTA	*PTA Design*	*PTA*PTA Design*
1	0	0	0
2	1	0	0
3	1	1	1

[10] For the handful of PTAs coded by Kucik (2012) but not by Dur, Baccini, and Elsig (forthcoming), we code their depth using the lowest possible value. But our results do not change in any significant way if we code these PTAs using the median *PTA Depth* value.

[11] http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx .

[12] We obtained these data directly from Jeffrey Kucik for his sample of PTAs.

number of escape options in form of safeguard (SG), anti-dumping (AD), and countervailing duties (CVD) provisions. PTAs with no such escape provisions are coded as 0, while those with all three escape features are coded as 3. This variable is included to capture the first conception of agreement flexibility (discussed earlier) and was coded using the data in Kucik (2012), but it was not part of his analysis in that paper.

To capture the second conception of agreement flexibility, we use the 16-point *PTA Restrictions* indicator from Kucik (2012). Given that most PTAs include some form of an escape provision, Kucik focused on the restrictions related to the use of these escape provisions with five possible restrictions related to each: 1) whether an investigation is required for authorization, 2) whether the use is eligible for dispute resolution, 3) whether the state must demonstrate an injury, 4) whether there is a limit on the size of the measure, and 5) whether there is a limit on the duration of the measure. Agreements with all five restrictions on each of these three possible escape features (SG, AD and CVD) are coded as *PTA Restrictions* =15. Finally, to capture the non-linear logic underlying the flexibility hypothesis, we also add a squared measure of Kucik's rigidity measure (*PTA Restrictions*^{2}).

It is important to note that our specification does not interact *PTA Escape* with the *PTA Restrictions* variables since the latter are nested with the former. If interacted, this nested structure would create nonsensical constitutive terms (e.g. *PTA Restrictions* measuring the effect of more escape provisions restrictions when there are no escape provisions within the PTA, or when *PTA Escape*=0). Thus, our flexibility/rigidity specification should be read additively: the number of escape provisions multiplied by the *PTA Escape* coefficient plus the number of restrictions multiplied by the *PTA Restrictions* coefficient plus the number of restrictions squared multiplied by the *PTA Restrictions*^{2} coefficient.

[Table 2 here]

Some descriptive statistics for all of these PTA design variables are presented in Table 2 Weighted in terms of dyad/year coverage (Table 2A), the average PTA is relatively shallow, plurilateral, includes a dispute settlement mechanism and at least one escape provision, but has relatively few escape restrictions. But unweighted (i.e. using the PTA unit of analysis in Table 2B), the average PTA shows greater depth, is bilateral, has more escape provisions with greater restrictions on their use. For some additional descriptive information about the PTA flexibility/rigidity features, Tables 3A and 3B present the frequency count for *PTA Escape* and *PTA Restrictions*.

[Tables 3A and 3B here]

Before proceeding, it is also useful to examine the bivariate correlations among the PTA design variables. Presented in Table 4 using both PTA dyad/years (Table 4A) and PTAs (Table 4B) as the unit of analysis, we find evidence to support the expected depth versus width tradeoff given the negative relationship between *PTA Depth* and *Plurilateral PTA*. Consistent with the depth versus rigidity tradeoff proposed by Johns (2014), there is a positive correlation between *PTA Depth* and *PTA Escape*, although deeper PTAs also have greater restrictions on the use of their more frequent escape provisions.[13] We also find evidence for a weaker width versus flexibility tradeoff based on the negative correlation between *Plurilateral PTA* and *PTA Escape*, although wider PTAs also have fewer restrictions on the use of their less frequent escape provisions.

[Table 4 here]

In the previous section, we considered two hypotheses about the effect of PTA flexibility/rigidity (in addition to the null hypothesis). As illustrated in Figure 1, the conventional

[13] This evidence is consistent with the results reported in Baccini, Dur and Elsig (2013).

wisdom and the flexibility hypothesis predict that unrestricted escape provisions should have a negative effect on trade, although the conventional wisdom expects a stronger negative impact than does the flexibility hypothesis. While both theories predict a negative sign for *PTA Escape*, we designate the larger negative coefficient expected by the conventional wisdom with double negative signs in Table 5.

[Table 5 here]

Both theories also expect that escape restrictions should have an initial positive effect on trade. Hence the conventional wisdom and the flexibility hypothesis predict a positive coefficient for *PTA Restrictions*, although this initial positive impact should be larger for the flexibility hypothesis (thus, its expectation for this variable is designated by the double positive signs in Table 5). Finally, the flexibility hypothesis expects a negative sign for *PTA Restrictions*^{2}, while the linear conventional wisdom expects *PTA Restrictions*^{2} to be statistically insignificant, or indistinguishable from zero (0).

Given that both propositions make similarly signed predictions for *PTA Escape* and *PTA Restrictions*, it is important to identify *in advance* how one can distinguish between these two theories for these two PTA design variables. *PTA Escape* is scaled 0-3 and *PTA Restrictions* is scaled 0-15 since 1 escape clause can have at most only 5 restrictions following Kucik's coding rules. Thus, in order for the negative impact of *PTA Escape* to be so large that *PTA Restrictions* can only partially offset this initial negative effect – as expected by the conventional wisdom - the absolute value of the negative *PTA Escape* coefficient must be 5 times greater than that of the positive *PTA Restrictions* coefficient. Conversely, if the absolute value of the negative *PTA Escape* coefficient is significantly less than 5 times greater than that of the positive *PTA Restrictions* coefficient, then the flexibility hypothesis's prediction that moderately restricted escape can have a net positive effect on trade could receive statistical support depending on the size of the negative *PTA Restrictions*^{2} coefficient.

Statistical Results

Having examined the PTA design variables and specified the variety of theoretical predictions, we enter them in our gravity model next to the *PTA* control variable (which will now pick up on *other* PTA effects beyond those related to the design variables) consistent with the full specification described earlier in equation (1). In model 2 in Table 1, we present the results when using the fullest possible sample (thus comparing among the PTA observations with the various design features, the PTA observations without the design features, and the non-PTA observations that also lack the design features) in order to reduce any sample selection bias.

Focusing on the flexibility/rigidity variables, one can see the statistically significant negative sign for *PTA Escape* and the statistically significant positive sign for *PTA Restrictions* expected by both the conventional wisdom and the flexibility hypothesis. As just discussed, the first key difference between these two propositions was the relative size of these two coefficients. In terms of this more specific prediction, the results in model 2 are consistent with the flexibility hypothesis as the absolute value of *PTA Escape* is not more than 5 times greater than the same for *PTA Restrictions*. Indeed, the two coefficients are approximately of equal size (-0.23 and 0.20), meaning that adding restrictions to the escape provisions *more than* offsets the negative effect of this flexibility feature (without any restrictions).

Further support for the flexibility hypothesis is seen in *PTA Restrictions*^{2}, which takes on a statistically significant negative sign (-0.007). This result means that adding more restrictions beyond a certain value has a negative impact on trade, consistent with the expectation that too much rigidity (just like too much flexibility) hurts trade cooperation. This relationship from model 2 is graphed in Figure 2A, plotting the marginal effect of *PTA Restrictions* (combining the positive effect of *PTA Restrictions* and negative effect of *PTA Restrictions*^{2}) on the y-axis with the possible values for *PTA Restrictions* (0-15) on the x-axis. The dashed lines correspond to the levels necessary to offset the negative effect of *PTA Escape* at each of its three non-zero values. From the results in model 2, four restrictions are enough to offset the negative impact associated with three escape provisions. One can also observe that while the negative coefficient for *PTA Restrictions^2* flattens the positive marginal effect of greater restrictions, the results in model 2 indicate that the most effective PTAs are those with 15 restrictions. More restrictions to the escape clauses could certainly be added (beyond just the 15 possible coded by Kucik 2012), but the results here suggest that any further restrictions would only have a negative impact on bilateral trade.

[Figures 2A, 2B, and 2C here]

While not of direct theoretical interest in this paper, it is important to note that a *Plurilateral PTA* appears to be more trade effective than a simply bilateral arrangement. Not surprisingly, PTAs with a dispute settlement mechanism are more trade effective than those without this feature, consistent with the results in Kono (2007). It is perhaps surprising to note that *PTA Depth* is statistically insignificant in model 2. If we drop the other design features, then *PTA Depth* would take on a statistically significant positive sign, consistent with the results in Dur, Baccini, and Elsig (forthcoming). So the insignificant result reported here may be due more to collinearity with other design variables than being a real indication that PTA depth does not matter much for its trade effectiveness. Finally, it is interesting to observe how the *PTA* control dummy takes on a statistically significant negative coefficient once this set of design variables are added next to it. This *PTA* sign flips (from model 1 to model 2) suggests that the statistically significant positive *PTA* coefficient in model 1 comes entirely from the attributes captured with this set of design features. *Other* PTA attributes, not modelled as a part of this design feature set, can be negatively associated with bilateral trade.

In model 3, we estimate this same specification but using a statistical sample that eliminates the observations associated with small values of dyadic trade. As noted by Goldstein, Rivers and Tomz (2007, 52), smaller trade flows may be problematic because "there is no way to distinguish between miniscule levels of trade, the absence of trade, and missing data." So we drop all observations associated with dyadic imports of less than $200,000 in 2009 US dollars. This reduces the size of the statistical sample by more than one-fourth, and eliminating the smallest trade flows attenuates the coefficient for all of the PTA design variables (except for *PTA Depth*, which gains some statistical significance with the expected positive sign). But even with the attenuated coefficients for the flexibility/rigidity variables, one can observe the same basic pattern as illustrated in Figure 2B. Unrestricted escape clauses have a negative effect on trade, but adding some restrictions to those escape clauses is more than enough to offset the initial negative effect (from the results in model 3, only three restrictions are necessary to offset the negative impact associated with the maximum number of three escape provisions). For this reduced sample, the most effective PTAs are those with no more than 13 escape restrictions; additional restrictions beyond this value only result in less overall trade effectiveness.

In model 4, we estimate our fully specified PTA design model using only the dyad/year observation with PTA coverage (this means that the *PTA* control variable will drop from the estimates). This greatly reduced sample eliminates any comparison to the observations that lack the design features because they are not covered by a PTA; thus, the comparison is now only between PTA observations with the design features and PTA observations lacking the design features. The obvious risk with this reduced sample is sample selection bias, but the results are remarkably similar to those already reported. Unrestricted escape has a negative effect on trade, but adding some restrictions more than offsets the negative effect associated with unconstrained flexibility. However, too many constraints are also associated with reduced trade. The flexibility/rigidity results from model 4 are graphed in Figure 2C.

To make one final demonstration of the robustness of these results, we now shift our unit of analysis from the directed dyad/year to the PTA itself, treating each preferential trading arrangement in our sample (N=301) as a single observation. Our trade effectiveness dependent variable comes from a full sample gravity model where we decompose the aggregated *PTA* variable used earlier into 301 separate preferential trading arrangements (no PTA design variables are entered into this model so the specification mirrors that in Table 1, model 1). From this gravity model (not presented for space considerations), we recover 301 PTA-specific coefficients, indicating how much each preferential trading arrangements in our sample can be associated with bilateral trade among its member-states. This research design parallels the one used by Kono (2007).

Due to differential dyad/year coverage in this gravity model, we obviously do not (and should not) have the same confidence in all of the 301 PTA-specific coefficients that now serve as our dependent variable. Thus, we estimate a weighted regression, using the PTA-specific t-statistic as the weight variable (with a larger value indicating greater statistical confidence in our estimated coefficient).[14] On the right-hand side, we include the six PTA design variables (*PTA Depth*, *Plurilateral PTA*, *PTA DSM*, *PTA Escape*, *PTA Restrictions*, and *PTA* $Restrictions^2$).

Our results are reported in Table 6. This estimate of trade effectiveness using the PTA unit of analysis returns design features coefficients that are largely similar to those in Table 1 for the directed dyad/year unit of analysis. *PTA Escape* is negatively associated with trade effectiveness, but this negative impact is more than offset by adding *PTA Restrictions*, although this effect turns negative beyond a certain value as shown by negative coefficient for *PTA* $Restrictions^2$. This non-linear relationship is graphed in Figure 3, which shows that six restrictions are sufficient to offset the negative impact of three escape clauses. But beyond 14 restrictions, PTA trade effectiveness moves in a negative direction.

3. Conclusion

To summarize, we find robust results for our flexibility/rigidity measures in a series of increasing restrictive statistical samples. The results are weakly consistent with the conventional

[14] For this weighting, we use the *aweight* command in Stata.

wisdom in that PTA flexibility in the form of escape provisions has a negative effect and that PTA rigidity in the form of escape restrictions has a positive effect on bilateral trade. But the results are more strongly consistent with the non-linear flexibility hypothesis. Adding *some* rigidity in the form of more escape restrictions has such a positive effect on trade that it more than offsets the negative effect associated with escape flexibility. But beyond a certain value, more escape restrictions begin to hinder PTA trade effectiveness. Thus, while too much PTA flexibility is clearly bad for trade, too much rigidity also appears to have a negative effect.

Having summarized our results, we now conclude with a discussion of their implications for both an academic and a policymaking audience. With regards to the former, our results have what we view as positive implications for the rational design of international institutions research program discussed in the introduction. Thus far in the RDII research program, scholars have primarily treated institutional design features - including but not limited to those related to flexibility - as dependent variables, seeking to explain when they are (or are not) selected. However, this focus only makes substantive sense if these design features actually have some causal effect on institutional performance. In this paper, we make this demonstration for PTA flexibility/rigidity features in terms of international trade. Furthermore, we show how these design features affect trade in both a positive and negative direction. It is thus rational for policymakers who want a particular trade outcome (presumably greater bilateral exchange) to select *for* escape provisions with *some* restrictions on their use.

This understanding leads directly to an important policy implication. Our results show that PTA trade effectiveness is maximized with a relatively large number of restrictions on the use of the escape provisions. Yet, the descriptive data show that most PTAs have only a few escape restrictions. Consequently, our results suggest that policymakers have *under-selected* for PTA escape restrictions. Stated differently, most PTAs would be even more trade effective with greater restrictions on the use of their escape clauses, although we recognize that trade promotion may not be the only goal of these bilateral/regional arrangements (even if it is a primary objective).

References

Anderson, James E. 1979. "A Theoretical Foundation for the Gravity Equation." *American Economic Review* 69 (1): 106-16.

Anderson, James E., and Eric van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review* 93 (1): 170-92.

Baccini, Leonardo, Andreas Dur, and Manfred Elsig. 2013. "Depth, Flexibility, and International Cooperation." Manuscript, London School of Economics.

Baier, Scott L., and Jeffrey H. Bergstrand. 2007. "Do Free Trade Agreements Actually Increase Members' International Trade?" *Journal of International Economics* 71 (March): 72-95.

Baldwin, Richard, and Daria Taglioni. 2006. "Gravity for Dummies and Dummies for Gravity Equations." *NBER Working Paper 12516*. Cambridge, MA: National Bureau of Economic Research.

Duffield, John S. 2003. "The Limits of 'Rational Design.'" *International Organization* 57 (Spring): 411-30.

Dunoff, Jeffrey L. 2010. "How Not to Think about Safeguards." In *Law and Economics of Contingent Protection in International Trade*, edited by Kyle W. Bagwell, George A Bermann, and Petros C. Mavroidis, pp. 401-11. Cambridge University Press.

Dur, Andreas, Leonardo Baccini, and Manfred Elsig. Forthcoming. "The Design of International Trade Agreements: Introducing a New Dataset." *Review of International Organizations*.

Goldstein, Judith L., Douglas Rivers, and Michael Tomz. 2007. "Institutions in International Relations: Understanding the Effects of the GATT and the WTO on World Trade." *International Organization* 61 (Winter): 37-67.

Gray, Julia, and Jonathan B. Slapin. 2012. "How effective are preferential trade agreements? Ask the experts." *The Review of International Organizations* 7(3): 309-333.

Head, Keith, and Thierry Mayer. 2014. "Gravity Equations: Workhorse, Toolkit, and Cookbook." *Handbook of International Economics* 4: 131-95.

Helfer, Laurence R. 2013. "Flexibility in International Agreements." In *Interdiscipinary Perspectives on International Law and International Relations*, edited by Jeffrey L. Dunoff and Mark A. Pollack, pp. 175-96. Cambridge University Press.

Jo, Hyeran, and Hyun Namgung. 2012. "Dispute Settlement Mechanisms in Preferential Trade Agreements Democracy, Boilerplates, and the Multilateral Trade Regime." *Journal of Conflict Resolution* 56 (6): 1041-1068.

Johns, Leslie. 2014. "Depth versus Rigidity in the Design of International Trade Agreements." *Journal of Theoretical Politics* 26: 468-495.

Jupille, Joseph, Brandy Jolliff, and Stefan Wojcik. 2013. "Regionalism in the World Polity." Manuscript, University of Colorado at Boulder.

Kono, Daniel Y. 2007. "Making Anarchy Work: International Legal Institutions and Trade Cooperation." *Journal of Politics* 69 (August): 746-59.

Koremenos, Barbara. 2005. "Contracting around International Uncertainty." *American Political Science Review* 99 (November): 549-65.

Koremenos, Barbara. 2001. "Loosening the Ties that Bind: A Learning Model of Agreement Flexibility." *International Organization* 55 (Spring): 289-325.

Koremenos, Barbara, Charles Lipson, and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55 (Autumn): 761-99.

Kucik, Jeffrey. 2012. "The Domestic Politics of Institutional Design: Producer Preferences over Trade Agreement Rules." *Economics and Politics* 24 (July) 95-118.

Kucik, Jeffrey, and Eric Reinhardt. 2008. "Does Flexibility Promote Cooperation? An Application to the Global Trade Regime." *International Organization* 62 (Summer): 477-505.

Mansfield, Edward D. 1998. "The Proliferation of Preferential Trading Arrangements." *Journal of Conflict Resolution* 42 (October): 523-43.

Mansfield, Edward D. and Eric Reinhardt. 2003. "Multilateral Determinants of Regionalism: The Effects of GATT/WTO on the Formation of Preferential Trading Arrangements." *International Organization* 57 (Fall): 829-62.

Martin, Lisa L., and Beth A. Simmons. 1998. "Theories and Empirical Studies of International Institutions." *International Organization* 52 (Autumn): 729-57.

Meyer, John, John Boli, George M. Thomas, and Francisco O. Ramirez. 1997. "World Society and the Nation-State." *American Journal of Sociology* 103 (1): 144-81.

Pelc, Krzysztof J. 2013. "The Cost of Wiggle-Room: Looking at the Welfare Effects of Flexibility in Tariff Rates at the WTO." *International Studies Quarterly* 57 (March) 91-102.

Pelc, Krzysztof J. 2009. "Seeking Escape: The Use of Escape Clauses in International Trade Agreements." *International Studies Quarterly* 53 (June): 349-68.

Rose, Andrew K. 2004. "Do We Really Know that the WTO Increases Trade?" *American Economic Review* 94 (March): 98-114.

Rosendorff, B. Peter, and Helen V. Milner. 2001. "The Optimal Design of International Trade Institutions: Uncertainty and Escape." *International Organization* 55 (Autumn): 829-57.

Tinbergen, Jan. 1962. *Shaping the World Economy: Suggestions for an International Economic Policy*. New York: Twentieth Century Fund.

Tomz, Michael, Judith L. Goldstein, and Douglas Rivers. 2007. "Do We Really Know that the WTO Increases Trade? Comment." *American Economic Review* 97 (December): 2005-18.

Wendt, Alexander. 2001. "Driving with the Rearview Mirror: On the Rational Science of Institutional Design." *International Organization* 55 (Autumn): 1019-49.

Table 1: Gravity Models of ln(*Imports*).

	1 Full sample	2 Full sample	3 *Imports*> $200,000	4 Only *PTA* dyad/years
GATT/WTO	0.21*** (0.03)	0.22*** (0.03)	0.11*** (0.02)	0.36*** (0.13)
PTA	0.43*** (0.03)	-0.59*** (0.10)	-0.32*** (0.08)	drops
PTA Depth		-0.002 (0.022)	0.03* (0.02)	-0.003 (0.042)
Plurilateral PTA		0.54*** (0.07)	0.33*** (0.05)	0.20* (0.12)
PTA DSM		0.29*** (0.07)	0.22*** (0.05)	-0.03 (0.13)
PTA Escape		-0.23** (0.09)	-0.11* (0.06)	-0.46** (0.23)
PTA Restrictions		0.20*** (0.04)	0.13*** (0.03)	0.27*** (0.10)
PTA Restrictions^2		-0.007*** (0.002)	-0.005*** (0.001)	-0.008** (0.004)
Constant	14.37*** (0.01)	14.37*** (0.01)	16.21*** (0.01)	15.44*** (0.23)
R^2	0.85	0.85	0.87	0.90
N	993,489	993,489	720,927	76,837

Cell entries are OLS coefficient with robust standard errors clustered on the dyad in parentheses.
Dyad, *Importer-Year* and *Exporter-Year* fixed effects are included, but not reported.
ln(*ProductGDP*) drops with the inclusion of *Importer-Year* and *Exporter-Year* fixed effects.
Distance drops with the inclusion of *Dyad* fixed effects.
Year fixed effects drop with the inclusion of *Importer-Year* and *Exporter-Year* fixed effects.
Statistical significance: *** p<.01, ** p<.05, *p<.10 (two-tailed).

Table 2: Descriptive Statistics for the PTA Design Variables

A. PTA Dyad/Year Unit of Analysis

	Mean	Std. Dev.	Min.	Max.
PTA Depth	0.88	1.26	0	7
Plurilateral PTA	0.90	0.30	0	1
PTA DSM	0.83	0.37	0	1
PTA Escape	1.33	0.75	0	3
PTA Restrictions	3.62	3.17	0	15
PTA Restrictions^2	23.15	45.40	0	225

B. PTA Unit of Analysis

	Mean	Std. Dev.	Min.	Max.
PTA Depth	2.15	1.73	0	7
Plurilateral PTA	0.16	0.37	0	1
PTA DSM	0.56	0.50	0	1
PTA Escape	2.00	0.94	0	3
PTA Restrictions	8.32	4.65	0	15
PTA Restrictions^2	90.85	73.28	0	225

Tables 3: Frequency Count for *PTA Escape* and *PTA Restrictions*

A. *PTA Escape*

PTA Escape =	# PTA dyad/years	# PTAs
0	5,766	32
1	46,982	35
2	16,737	132
3	7,352	102
	N=76,837	N=301

B. *PTA Restrictions*

PTA Restrictions =	# PTA dyad/years	# PTAs
0	5,766	32
1	66	3
2	35,640	19
3	7,283	7
4	11,124	9
5	7,269	5
6	1,398	3
7	297	5
8	2,039	101
9	293	11
10	147	5
11	94	3
12	1,247	7
13	1,777	36
14	1,854	38
15	543	17
	N=76,837	N=301

Table 4: Bivariate Correlations among the PTA Design Variables

A. PTA Dyad/Year Unit of Analysis

	PTA Depth	*Plurilateral PTA*	*PTA DSM*	*PTA Escape*	*PTA Restrictions*	*PTA Restrictions^2*
PTA Depth	1.00					
Plurilateral PTA	-0.33	1.00				
PTA DSM	0.11	0.11	1.00			
PTA Escape	0.59	-0.25	0.11	1.00		
PTA Restrictions	0.50	-0.39	0.06	0.85	1.00	
PTA Restrictions^2	0.43	-0.40	0.04	0.73	0.96	1.00

B. PTA Unit of Analysis

	PTA Depth	*Plurilateral PTA*	*PTA DSM*	*PTA Escape*	*PTA Restrictions*	*PTA Restrictions^2*
PTA Depth	1.00					
Plurilateral PTA	-0.09	1.00				
PTA DSM	0.26	0.14	1.00			
PTA Escape	0.31	-0.30	0.10	1.00		
PTA Restrictions	0.35	-0.29	0.16	0.96	1.00	
PTA Restrictions^2	0.36	-0.23	0.28	0.88	0.96	1.00

Table 5: Expected Signs for the PTA Flexibility/Rigidity Indicators

	PTA Escape	*PTA Restrictions*	*PTA Restrictions^2*
Conventional Wisdom	- -	+	0
Flexibility Hypothesis	-	+ +	-

Table 6: Estimate of Trade Effectiveness
Using the PTA Unit of Analysis

PTA Depth	-0.17*** (0.04)
Plurilateral PTA	1.07*** (0.19)
PTA DSM	0.24 (0.17)
PTA Escape	-0.63** (0.29)
PTA Restrictions	0.38*** (0.10)
PTA Restrictions ^{2}	-0.013*** (0.004)
Constant	-0.61** (0.24)
R^2	0.17
N	301

Dependent variable is the PTA-specific coefficient from the gravity model of trade.
Observations weighted by the PTA-specific t-statistic from the gravity model of trade.
Statistical significance: *** p<.01, ** p<.05, and *p<.10 (two-tailed).

Figure 1: Hypotheses about Trade Agreement Flexibility/Rigidity

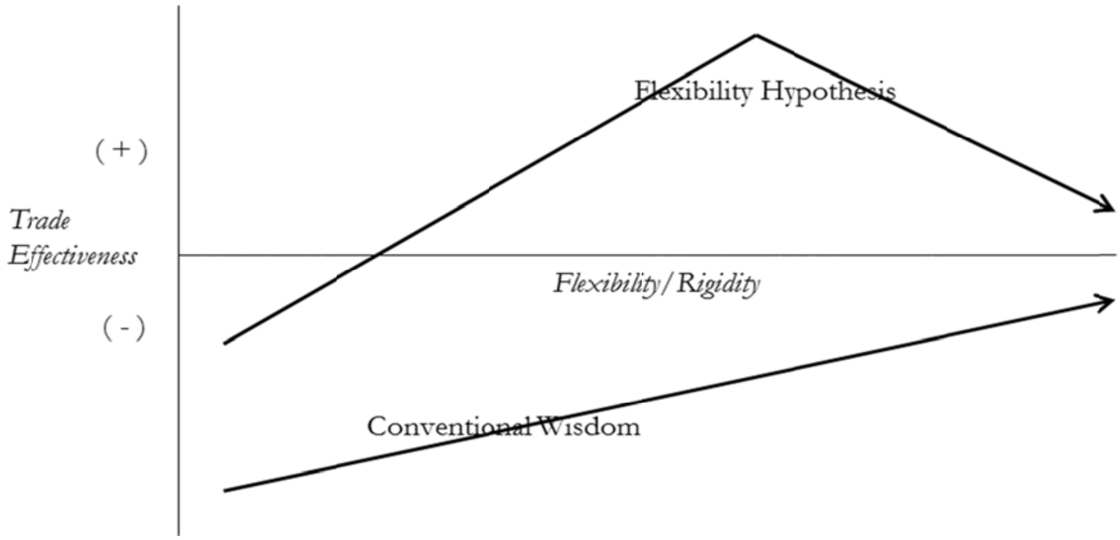

Figure 2A: Marginal Effect of *PTA Restrictions* from Model 2

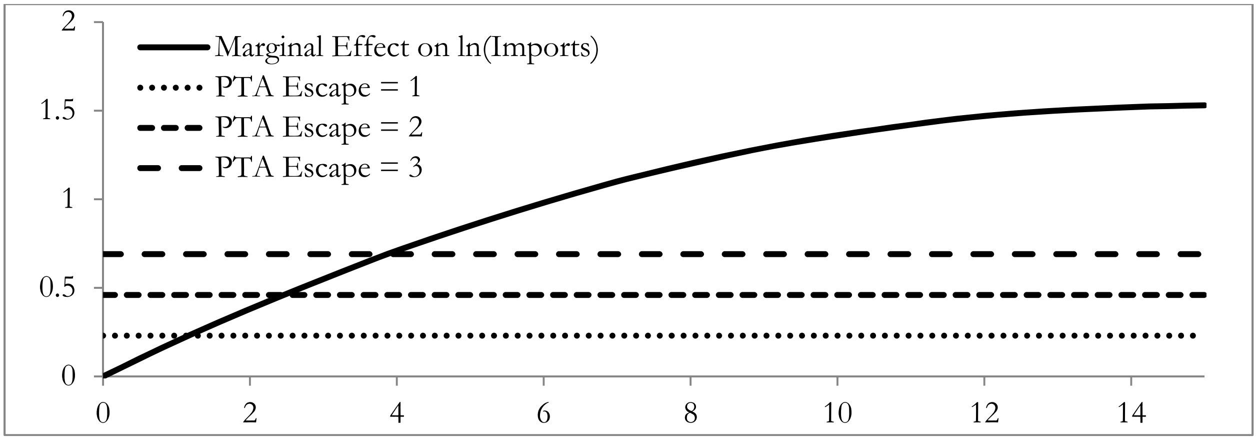

Figure: 2B: Marginal Effect of *PTA Restrictions* from Model 3

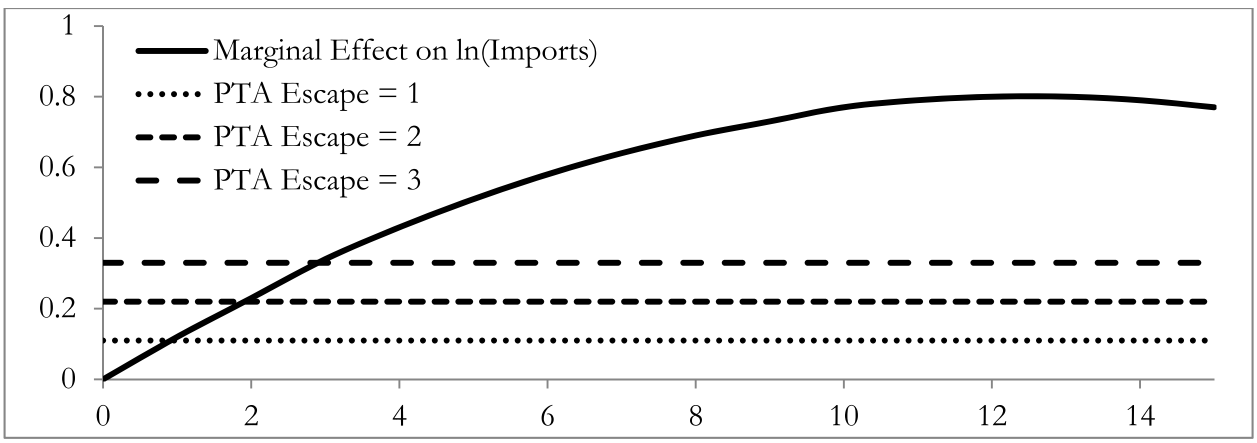

Figure: 2C: Marginal Effect of *PTA Restrictions* from Model 4

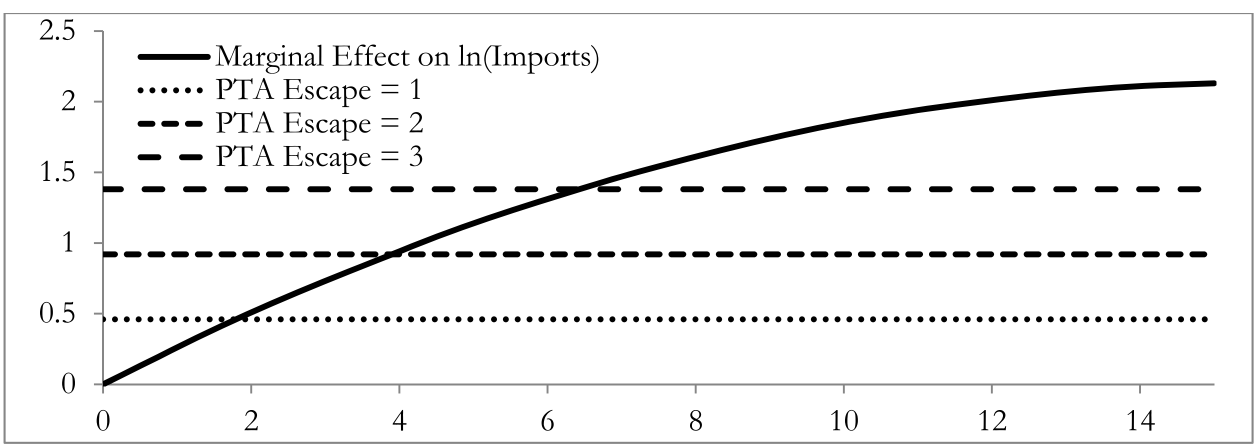

Figure 3: Marginal Effect of *PTA Restrictions* from Table 6