Depth, Flexibility and International Cooperation: The Politics of Trade Agreement Design

Leonardo Baccini, LSE

Andreas Dür, University of Salzburg

Manfred Elsig, University of Bern

Paper prepared for presentation at the ISA Annual Conference in San Francisco, April 3-6, 2013

Abstract

What explains the degree of flexibility included in international agreements? Focusing on preferential trade agreements, we argue that the depth of an agreement is a major determinant of its flexibility. Deep agreements create an incentive for states to add more flexibility of both a short-term and a long-term type. Importantly, therefore, different aspects of institutional design are not independent. Moreover, we submit that as states introduce flexibility, they will make long-term flexibility more difficult to use. That is, we expect more flexibility and more rigidity to go hand in hand. An original dataset on the design of 587 trade agreements signed between 1945 and 2009 allows us to test our arguments. Descriptive evidence, and ordered probit, tobit and instrumental variable models all support the theoretical expectations. The paper contributes to the literatures on the design of international institutions, international cooperation and the political economy of international trade.

Key Words: design of international institutions, trade agreements, depth, flexibility, lobbying.

Introduction

International institutions strongly vary in their design. In particular, over time states have increased the depth and legalization of international agreements. In parallel to this "move to law" (Goldstein et al. 2000), states have also included a large number of flexibility measures such as escape clauses in international agreements, which allow them to react to changing domestic conditions or international challenges (Koremenos et al. 2001; Helfer 2013). Is this move towards flexibility a response to the greater depth of agreements? That is, does greater depth require more flexibility in international agreements?

We focus on international economic cooperation, and in particular on bilateral and regional trade agreements, in responding to this question. Since 1947, states have concluded well beyond 700 trade agreements (Dür et al. 2013). Especially the last twenty years have seen a rapid increase in the number of such agreements. These preferential trade agreements (PTAs) are an attractive focus for our study as they are a worldwide phenomenon and as we observe large variation across agreements in terms of design. In focusing on PTAs, our study also nicely complements the existing literature on the design of international trade institutions that so far has mainly focused on the World Trade Organization (WTO) and its forerunner the General Agreement on Tariffs and Trade (GATT) (Rosendorff 2005; Rosendorff and Milner 2001; Kucik and Reinhardt 2008; Pelc 2009).

Our argument that depth indeed creates demand for more flexibility is in line with existing research. Relying on a formal model, for example, Johns (2012) argues that "deep agreements will be flexible, while shallow agreements will be rigid". In at least two respects, however, we go beyond this literature. First, we distinguish between two types of flexibility, namely short (tariff transition periods) and long-term flexibility (escape clauses, anti-dumping and anti-subsidies provisions). We argue that with increasing depth, both types of flexibility will increase. Second, and more interestingly, we submit that states, while granting more flexibility, at the same time also increasingly restrict the use of some flexibility measures. Flexibility and rigidity hence may go hand in hand.

We test our expectations on a novel dataset of 587 trade agreements signed between 1945 and 2009 (Dür et al. 2013). This is by far the most comprehensive dataset on the design of trade agreements in terms of both agreements covered and detail of the coding. Our coding of the texts of these agreements allows us to establish measures of depth and flexibility for each of them. Descriptive evidence, and the results from ordered probit, tobit and instrumental variable models all offer clear-cut support for our theoretical expectations. The findings are highly robust to changes in estimation method and the operationalization of variables.

In studying the design of international institutions, the paper speaks to the debate be-

tween realists and neoliberal institutionalists concerning the role that institutions play in international relations. Realists submit that state interests and the distribution of power dictate institutional design (Mearsheimer 1994). Neoliberal institutionalists propose that institutional design is mainly a result of the functions the international institution is expected to fulfil (Stein 1982). We show that neither power nor function dictate form; states have leeway in designing institutions and they use this leeway to create institutions that maximize support and minimize opposition among domestic interests.

Our study's more specific contribution to the literature on the design of international institutions (Koremenos et al., 2001) is to show that different design features are interdependent.[1] Treating an aspect of institutional design in isolation from other aspects seems problematic in view of this finding. Moreover, the paper speaks to the literatures on trade and the political economy of PTAs (Mansfield et al., 2002; Chase, 2005; Baccini and Dür, 2012). In this regard, our results are suggestive of the role of societal interests in shaping PTAs in general, and their design in particular.

In the following sections, we first discuss the literature on flexibility (Section 1) and then develop our argument (Section 2). Section 3 introduces our approach to empirically testing the resulting theoretical expectations and also features some descriptive evidence. Section four, finally, presents the empirical findings and robustness checks.

Flexibility: What Do We Know?

Flexibility provisions are generally understood as devices included in an agreement that allow states to anticipate and respond to domestic contingencies or to adjust their policies for other purposes without violating the terms of an agreement. Flexibility clauses then provide for legally accepted opt-outs without leading to a *de jure* breach of an agreement.

A rich literature studies flexibility in international cooperation (see Koremenos et al. 2001; Helfer 2013). The flexibility instruments discussed in this literature include membership rules and voting procedures (Goodman and Jinks 2004), soft as opposed to hard law provisions (Shaffer and Pollack 2010), degrees of delegation (Alvarez 2005), exit options (Helfer 2005), duration and renegotiation clauses (Koremenos 2005), reservations (Neumayer 2007), escape clauses (Rosendorff and Milner 2001), and withdrawal clauses (Koremenos and Nau 2010). Koremenos et al. (2001: 773) classify these instruments into two types, namely transformative and adaptive flexibility. The former allows for significantly changing the terms of obligations (e.g. re-negotiations or exit), the latter makes possible an individual escape as a temporal exception which however does not

[1] Haftel (2013) makes a similar point when relating the scope of agreements to variation in the institutionalization and design of dispute settlement mechanisms in twenty-eight regional economic organizations. Johns (2012) relates depth and rigidity, but does not attempt to empirically test the argument.

question the stability of an agreement.

A significant proportion of the flexibility literature focuses on the WTO and the GATT. Pelc (2009) shows how GATT escape clauses have changed over time, with different flexibility measures serving different purposes. In another contribution, he provides evidence that the degree of flexibility (defined as the amount of overhang in tariff levels) is a function of how countries can use currency depreciation to react to crises and whether they have access to trade remedies (Pelc 2010). More options to react to future challenges will lead to less flexibility. This work suggests that flexibility depends on the perceived possibilities to use these in the future and that there are important interaction effects between flexibility tools. The latter is in line with the more general work by Koremenos (2005) who focused on how various flexibility tools may act as substitutes (see also Helfer 2013). Little work has been done on flexibility in PTAs, with the notable exception of Kucik (2012) who focuses on variation in flexibility provisions in relation to trade remedies. He argues that the political struggle between exporters and import-competitors influences the flexibility of PTAs.

While this literature addresses a large number of important questions, so far our understanding of why we see variation across agreements in terms of flexibility is limited. Some studies suggest that issue characteristics are important for the degree of flexibility granted in agreements; others focus on member state characteristics. We argue, by contrast, that other features of the design of an agreement, in particular the depth of the agreement, are crucial in determining the flexibility of an agreement. Below, we develop this argument.

Theory and Hypotheses

Depth in PTAs

We assume that states cooperate with each other in the trade field with the aim of improving foreign market access for their exporters (Dür 2010, Elsig and Dupont 2012, Manger 2009). They sign trade agreements that exchange domestic for foreign market access. If this is done in a preferential manner, the resulting trade agreement is likely to boost exports from the member countries while hurting exporters from third countries.

Trade agreements strongly vary with respect to their depth, that is, the extent to which they liberalize trade. While the average tariff cuts are an important determinant of depth, they are not the only one. PTAs can also contribute to liberalizing trade relations between states by including concessions in opening some of the services sectors or by allowing foreign companies to bid for tenders in government procurement. Their impact on trade flows are sometimes higher than through cuts in tariffs (Dür et al. 2013), as they

provide new market access (extensive margins) in addition to improved market access (intensive margins) for exporters. Similarly, the protection of foreign direct investment can substantially enhance market access for exporters. Some agreements also foresee improved treatment of foreign direct investments over and above the protection of investments against arbitrary provisions (see also World Trade Organization 2011).

Other behind-the-border obstacles that a PTA may remove are burdensome technical standards, sanitary or phytosanitary measures, inadequate protection of intellectual property rights, and competition rules that discriminate against foreign traders. Some PTAs, for example, foresee the mutual recognition of international product standards or the harmonization of technical regulations (Piermartini and Budetta 2009). This reduces transaction costs and increases market integration. The strengthening of intellectual property rights is important for exporters to pursue a long-term market penetration strategy. Absence of patent or trade mark protection directly translates into less exports and less technology transfer in the case of investments directed at foreign markets (Maskus and Penubarti 1995). Finally, cooperation on competition policy is important to address unfair competition by state enterprises and private firms. Similar to intellectual property rights it creates a regulatory environment that strengthens the rule of law. This provides important incentives for exporting firms to decide to pursue active market expansion in the PTA partner's home economy.

Designing measures to level the field between domestic and foreign firms clearly positively impact on exporters' market shares. Therefore, by signing deep trade agreements, governments can offer better market access to exporters, which is reflected in positive trade flow effects (Baier and Bergstrand 2007; Dür et al. 2013).

Types of flexibility in PTAs

At the same time as governments strive to improve market access for exporters through reciprocal deals with foreign governments, we expect import-competitors to oppose these agreements or at least to ask for protection for their sectors. For that purpose, PTAs can include a set of escape clauses or other flexibility devices. These flexibility tools most likely are of an adaptive nature in the sense that they allow for the future temporal protection of the economy (or sectors thereof). These mechanisms can be used in times of economic shocks and sudden changes in the economic environment, such as economic downturns. Therefore, PTAs foresee a set of safety valves to allow for a temporal legal breach in case imports surge as a result of increased trade flows between PTA partners (see also Kucik 2012). These include balance of payments (BoP) exceptions, general or specific safeguard provisions, as well as possibilities to use anti-dumping and countervailing duty provisions to address dumping by foreign firms or subsidies by foreign governments. We call these

types of flexibility measures "long-term flexibility".

While import-competitors benefit from long-term flexibility, exporters that value not only good access to foreign markets, but also the stability of trade, are likely to oppose such flexibility (Kucik 2012:98). If foreign governments use flexibility mechanisms without restraint, the benefits of market access may be nullified. Even if exporters do not have the clout to fully avoid the inclusion of long-term flexibility, they thus will try to restrict its use. They will ask governments to bind their own and the other governments' hands. Various possibilities exist to make flexibility measures more rigid. In terms of safeguards and antidumping, states may limit the duration of an anti-dumping duty, restrict the upper level of the trade remedy imposed, or make reference to the GATT/WTO legal framework that prescribes a number of procedural and substantive obligations that control against abuse. In terms of subsidies, treaties may foresee governments to decrease or eliminate subsidies or demand an active cooperation among authorities to address competition-related negative spill-overs. We call these measures "rigidity on flexibility".

We expect import-competing groups, if they cannot torpedo the agreement in advance (an example for this would be the Switzerland-US PTA that was stopped after an exploratory study and concerns by the farm lobby), to also pursue temporal protection within the limits of WTO law. Governments design flexibility in PTAs to allow domestic industries that are threatened by more competition through imports additional time to prepare for liberalization. This is a short-term type of flexibility which is negotiated *ex ante*. Governments can provide flexibility to these interest groups during the implementation period until full concessions kick in. This is usually done by negotiating long transition periods in terms of tariff liberalization. This is a direct way to give some short-term flexibility to certain firms to prepare for full liberalization. The GATT/WTO defines an upper-level of accepted years for implementation. Many PTAs foresee further negotiations over faster implementation in case competition is less acute than anticipated. Therefore this type of flexibility, which we call "short-term flexibility", does not need to be fully used by governments.

The relationship between depth and flexibility

What explains the design of these flexibility measures and the overall balance between depth and flexibility? We assume that while governments act as catalysts for exporters' interests, they cannot ignore import-competing industries' concerns related to negotiating PTAs. This is so because governments that want to remain in office are sensitive to the preferences of politically mobilized interest groups (Milner 1988; Gilligan 1997). The underlying assumption is that governments depend on business support for re-election and the formulation and implementation of policies (see, for example, Wright 1996; Hall

and Deardorff 2006). In a re-election effort (or the efforts of an autocratic government to stay in power), supportive business can provide campaign financing or invest resources in influencing public opinion. Supportive business actors can also be expected to offer information on the technical and political feasibility of a policy that is important for the viability of a proposal.

Governments do not adhere to a "winner-takes-all" logic, however. Aiming to maximize the chances of staying in power, to the extent possible, they try to satisfy all sides in a policy debate, in this case import-competitors and exporters. They can do so by including or excluding certain provisions from trade agreements. Trading powers such as the European Union, Japan and South Korea protect their agricultural sectors when agreeing to negotiate a trade agreement. Similarly, the US insists on the inclusion of environmental chapters that benefit import-competitors in its trade agreements. Governments thus do not face a dichotomous choice between signing or not signing a PTA, but a choice among many options ranging from no agreement to a very deep and rigid agreement, with agreements with varying degrees of depth and flexibility in between.

We make two related arguments as to how governments address exporters and import-competitors' using long-term and short-term flexibility. First, in relation to long-term flexibility, we argue that the expectation is that governments will design more flexibility in deeper agreements. Governments, however, will seek an optimal degree of flexibility that allows for temporal breach and adaptation if necessary, but can be restricted through a set of rigidity tools. They face a classical time-inconsistency problem in designing escape clauses. They know they will be tempted to use such clauses in the future (as doing so may be important for staying in office). But governments also assume that unlimited recourse to these opt-outs endangers the overall benefits of the agreement (and its stability). Therefore, they have a strong incentive ex-ante to define the upper limits and procedural constraints to control their application. Governments need to balance both interests by important import-competing constituencies to use flexibility and exporters' worries about being arbitrarily hit by a PTA partner's long term flexibility. If we combine both the government's interest to explore the possibilities of deep agreements and the need for balanced flexibility, this allows to predict the trade-offs governments settle for in relation to escape clauses. The trade-off between depth and rigidity exists (Johns 2012), but the use of flexibility instruments is conditioned by constraints which in turn decreases overall flexibility. From the above discussion, we derive two hypotheses about the relationship between depth and flexibility as well as flexibility and rigidity on flexibility tools in relation to classical trade escape clauses.

Hypothesis 1 *As the depth of an agreement increases, long-term flexibility also increases.*

Hypothesis 2 *As long-term flexibility increases, we see more rigidity on flexibility.*

Second, governments can rely on multiple forms of providing flexibility which serve different purposes. Next to long-term flexibility, governments may also be forced to offer short-term flexibility. With increasing depth of an agreement, import-competing groups will lobby more for getting temporal opt-outs in form of longer transition periods. The prospect of a deep agreement will intensify endogenous lobbying by import-competing firms. The existence of classical long-term flexibility is not perceived to be sufficient. Import-competing groups cannot wait until some unforeseen developments in the future threaten the competitiveness of the sector or the firms. Governments will therefore be tempted to provide short-term flexibility which increases with the ambition of market integration. In the case of the EU-South Korean trade negotiations, the European Commission not only provided for a special safeguard to the European car industry, but also offered the longest possible transition period for lowering tariffs on key car and truck imports (Elsig and Dupont 2012). In this case, it is not the time-inconsistency that drives government behavior towards designing this type of flexibility, but the necessity to address the most vocal criticism towards a deep PTA in the short run.

Hypothesis 3 *With increasing depth of an agreement, we witness more short-term flexibility.*

Operationalization of variables and descriptive evidence

For the empirical analysis, we rely on an original dataset on the design of 587 PTAs signed between 1945 and 2009. We coded agreements for a total of ten broad sectors of cooperation that may be included in PTAs, encompassing market access, services, investments, intellectual property rights, competition, public procurement, standards, trade remedies, non-trade issues, and dispute settlement. For each of these sectors, we have coded a significant number of items, meaning that we have about 100 data points for each agreement. The coding has been carried out manually, with the results cross-checked against existing datasets that partially overlap with ours (for more detail, see Dür et al. 2013). This dataset provides us with a unique opportunity to measure depth and flexibility.

Depth refers to the amount of liberalization envisaged by an agreement. We use two different measures of depth below. On the one hand, we created a simple additive index of 48 items in our dataset ($Depth(index)$). We only included items in the index that theoretically are related with depth and weighted each item equally (the annex contains a list of all of these items). On the other hand, we used exploratory factor analysis, a technique that is used to reveal unobserved factors underlying a set of variables, to arrive

at a measure of depth ($Depth(FA)$). Factor analysis uses the correlations across items to determine the weight that an item should get in a measure of depth. Because our dataset includes dichotomous measures of the design of PTAs, we use tetrachoric factor analysis (Revelle 2012). We ran the analysis on the same variables that also enter the additive index, extracted three factors, rotated them (using varimax rotation), and then calculated scores for the first factor while assuming that depth is a continuous variable. Importantly, the two measures are highly correlated with each other (r=0.72). Figure 1 shows this correlation graphically. For a couple of agreements, including several EU agreements with Central and Eastern European countries, $Depth(FA)$ seems to underestimate depth.

Figure 1 About Here

Hypothesis 1 refers to a variable called long-term flexibility. We measure this variable using a simple additive index of the presence or absence of four provisions in PTAs: a provision allowing for the suspension of tariff cuts in case of balance of payments problems, a general safeguard provision, a provision allowing for the imposition of countervailing duties, and a provision allowing for the imposition of anti-dumping duties. In the absence of these provisions, a country that suspends its tariff cuts or imposes antidumping and countervailing duties for goods covered by the agreement is in breach of the agreement. These four provisions thus serve as escape clauses. The index can range from 0 to 4. In our dataset, 73 agreements score 0 and 260 agreements score 4 on this variable. Over time, we see an increasing number of agreements including many of these escape provisions.

Following Hypothesis 1, this variable should be positively related to depth. To examine this proposition, we recoded the two depth variables into ordinal variables that range from 0 (very shallow) to 4 (very deep) and then calculated the mean number of flexibility provisions for each level of depth. The resulting graph (Figure 2) shows a strong positive relationship between the two variables when using $Depth(index)$, but the relationship is not linear when using $Depth(FA)$.

Figure 2 About Here

Hypothesis 2 refers to rigidity on flexibility. We operationalize this variable using six items that capture rigidity with respect to the imposition of antidumping duties, the provision of subsidies, and the use of the safeguard provisions. These capture provisions that impose WTO rules on the use of the antidumping instrument, the safeguard provision and the provision of subsidies, stipulate that the safeguard provision is only valid during a transition period, create a common policy on subsidies, and stipulate a minimum dumping margin. The variable potentially ranges from 0 to 6, depending on the number of provisions included in a trade agreement that restrict the imposition of antidumping

duties and the use of subsidies and the safeguard provisions. In practice, the variable ranges from 0 (195 agreements) to 5 (4 agreements).

Figure 3 offers a first test of our argument concerning a relationship between long-term flexibility and rigidity on flexibility. The relationship between the two variables is positive as expected following Hypothesis 2. As long-term flexibility increases, countries make their use more difficult.

Figure 3 About Here

Finally, in Hypothesis 3 we stipulate a relationship between depth and short-term flexibility. We operationalize short-term flexibility relying on the maximum number of years that countries are allowed to achieve the liberalization of tariffs envisaged in the agreement. The longer the transition period, the more flexibility exists for import-competing groups to adjust to increased competition. Phase out periods for tariff liberalization range between 0 years (all tariffs are liberalized at the date of entry into force of an agreement) and 20 years (usually for a selected number of sensitive products). We took the arithmetic average for the member countries of a PTA if the transition periods vary across countries.

In Figure 4 we show the relationship between depth and short-term flexibility. As expected in Hypothesis 3, we see a positive relationship between the two variables. As depth increases, also short-term flexibility increases.

Figure 4 About Here

Empirical Models

By choosing PTA as unit of analysis, we depart from previous studies that use dyad-year (Mansfield et al., 2002) or PTA-country-year (Kucik 2012). The reasons for our choice are two-fold. First, the design of a PTA does not vary across member countries in our dataset. Thus, if we use dyad-year or PTA-country-year as units of analysis, we would duplicate the value of our dependent (and independent) variables in plurilateral agreements. Second, the design of PTAs does not vary over time in our dataset with the exception of the EU. Thus, there is no loss of information by dropping the time dimension. In other words, using PTA as unit of analysis is a conservative choice to avoid inflating the number of observations.

Going beyond the descriptive evidence, we estimate regression models that first explain the three variables *Long-term Flexibility*, *Rigidity on Flexibility* and *Short-term Flexibility* separately. This analysis is implemented using ordered probit and tobit models. Next, we address the causal mechanisms highlighted by our theory. Specifically, we rely on an

instrumental variable analysis, which takes into account the fact that *Long-term Flexibility* and *Short-term Flexibility* are endogenous to *Depth*, and *Rigidity on Flexibility* is endogenous to *Long-term Flexibility*.

The empirical estimation presents several challenges. First, it may be argued that states decide on the depth and flexibility of international agreements, simultaneously. That is, it may be that not only depth influence flexibility, but also flexibility depth. To account for this possible endogeneity, we need to find suitable variables to instrument *Depth* and *Long-term Flexibility*. Below we describe our identification strategy, showing the theoretical plausibility of our instruments, their strengths, and how they meet the exogeneity condition. Second, since our main variables are ordinal and interval, we cannot rely on a classic IVREG estimation, which would produce biased and inconsistent coefficients. Thus, we use conditional recursive mixed process estimators developed by Roodman (2011). Mixed process implies that different equations can have different kinds of dependent variables (response types), i.e. interval-censored and ordered probit. That grants the flexibility to regress an interval variable on an endogenous ordinal one, for instance. Recursive implies that we are only able to fit sets of equations with clearly defined stages, whereas we are unable to estimate simultaneous causation. In particular, we rely on a limited-information (LIML) estimator, in which only the final stage's coefficients are structural, while the first stage are reduced-form parameters.[2]

Formally, we estimate the following three sets of equations:

$$Long-termFlexibility_i = \alpha_1 + \beta_1(Depth_i = Z1_i) + \beta_2 X_i + \epsilon_1. \quad (1)$$

$$RigidityonFlexibility_i = \alpha_2 + \beta_3(Long-termFlexibility_i = Z2_i) + \beta_4 X_i + \epsilon_2. \quad (2)$$

$$Short-termFlexibility_i = \alpha_3 + \beta_5(Depth_i = Z3_i) + \beta_6 X_i + \epsilon_3. \quad (3)$$

where *Long-term Flexibility*, *Rigidity on Flexibility*, and *Short-term Flexibility* are dependent variables. *Depth* is the main independent variable in equations 1 and 3, whereas *Long-term Flexibility* is the main independent variable in equation 2. X_i are vectors of control variables, $Z1_i$, $Z2_i$, and $Z3_i$ are instruments. β_1, β_2, ..., β_6 are the coefficients, α_1, α_{12}, and α_2 are constants, and ϵ_1, ϵ_2, and ϵ_2 are the error terms.

[2]We use the command 'cmp' in Stata 12. The estimation relies on the Geweke-Hajivassiliou-Keane algorithm for estimating cumulative normal densities above dimension (command 'ghk2' in Stata 12). Differently from GLLAMM, 'cmp' uses simulation. That is, it makes many draws from the hypothesized normal distributions of the effects, computing the implied likelihood for each set of draws before averaging (Train 2009; Greene 2011, chap 15).

We note that equation 1 is a tobit in the first stage and an ordered probit in the second stage. Equation 2 is an ordered probit in both first and second stage. Equation 3 is a tobit model in both first and second stage. We have already discussed the design variables in the previous section. Below we discuss the instruments and the operationalization of the control variables.

Identification Strategy

To correctly identify our models, we need to find an instrument that is a good predictor of *Depth*, but that is unrelated with the error term of *Long-term Flexibility* and *Short-term Flexibility*. That is, a good instrument should be theoretically related to one of the two dependent variables, but should not explain the other one *once we include the control variables that account for alternative causal paths*.

To instrument *Depth* we use the depth of GATT/WTO agreements signed by member countries in the various negotiation rounds (*GATT/WTO Depth*). Since the GATT/WTO is the largest and most important trade agreement, it is safe to assume that provisions bargained and included in the GATT/WTO treaties inspired the design of bilateral and plurilateral agreements. For instance, the signature of agreements concerning intellectual property rights and trade-related investment measures in the WTO triggered the inclusion of similar provisions in PTAs. Equally, the services provisions included in PTAs are very similar to those of the General Agreements on Trade in Services (World Trade Organization 2011: 134). To empirically measure the depth of GATT/WTO agreements, we use the same operationalization as for $Depth(index)$. Importantly, this operationalization only includes items related to the depth of agreements; this ensures that *GATT/WTO Depth* does not have a direct impact on the flexibility of PTAs.

We also instrument the depth of PTAs using the score of the deepest EU PTA ($EUDepth$). The reason for this choice is that the EU provides "the standard model for regional integration" (Börzel and Risse 2012, 196). In other words, the EU could serve a similar purpose than the one outlined above in respect to the GATT/WTO. The correlation between $WTODepth$ and the error term is 0.1, whereas the correlation between $EUDepth$ and the error term is 0.05.[3]

To instrument *Long-term Flexibility*, we rely on a variable that captures the level of transparency of PTA member countries, since earlier studies show that transparency affects the supply side of flexibility (Svolik 2006; Baccini 2010). The reasoning is that domestic political changes, adjustment costs, and distributional problems constitute private

[3]If we run classic instrumental variable regressions in equation 1 and 3, the Cragg-Donald Wald F test of the first stage is 66.04 and 63.32, respectively. Thus, there is no concern that our instruments are weak. Moreover, results from the Sargan-Hansen test show no concern of overidentification, which means that the instruments are valid.

information. Countries then have an incentive to misrepresent their private information in order to free-ride on cooperation. Since each country has an incentive to overstate their costs of compliance at any time, in the absence of transparency trade partners cannot distinguish between the use of an escape clause due to serious difficulties or due to opportunistic behaviour. This follows naturally from Bayesian updating, as the sources of any given defection can be seen as coming from either forced emergency measures or opportunism, and is in line with previous studies in the field (Svolik 2006). For instance, we note that the US-Vietnam PTA (2000) is less flexible than it would be expected from its level of depth. Vietnam's low level of transparency might account for this 'missing' flexibility.

Operationalizing transparency is difficult. We rely on the amount of information that states are willing (and able) to release. Since measures of flexibility are invoked due to economic difficulties, the capacity of communicating on the status of a country's economy seems to us a more direct way of capturing transparency than the sole presence of a democratic regime. Concretely, we use the measure of transparency developed by Hollyer, Rosendorff, and Vreeland (2011) who look at the availability (or absence) of policy-relevant data in the World Development Indicators database (World Bank 2012). Because the resulting variable ($Transparency$) has limited variation, we use a linear regressor and the quadratic regressor, which are both statistically significant at the conventional level.

In theory, $Transparency$ could be endogenous to $Depth$. However, including a democracy score on the right-hand side should account for the impact of domestic institutions. $Transparency$ thus should not be correlated with the error term. For instance, Svolik (2006) conceptualizes transparency with the presence of a democratic regime in general, and elections, in particular. In other words, the residuals of $Depth$ should be uncorrelated with $Transparency$ and $Transparency^2$, granting the validity of the exclusion restriction. Indeed, Figure 8 in the Appendix shows that there is no correlation between $Transparency$ and the error term of *Rigidity on Flexibility*, i.e. $\rho = -0.15$ for $Transparency$ and $Transparency^2$.[4]

Control Variables

We include several control variables that allow us to deal with potentially confounding factors. Since our unit of analysis is PTA, we take the mean of each variable across member countries for continuous variables.[5] In doing so, we acknowledge that monadic and dyadic variables lose variation. That is the price that we pay for having a clean research design in our variables of interest.

[4]The error term are the residuals of the extended model shown in the next section (Model 8).

[5]Our results are similar if we use the median or minimum value across member countries.

In terms of economic variables, we use GDP per capita ($GDPpc$) and total GDP (GDP) to capture the income level and economic importance of a country. Larger and richer countries might be able to negotiate agreements with a higher degree of depth and less flexibility compared to poorer or smaller countries. Moreover, we add $Trade$, which is the mean of the value of exports and import across all the members of a PTA. As is practice in the empirical trade literature, we take the log of this value. We expect the amount of trade to positively influence the demand for agreements with both a high degree of depth and flexibility.

Previous studies show that domestic institutions are important determinants of the design of international institutions and international organizations (Mansfield et al. 2008). We thus include a democracy score ($Regime$), which comes from Polity IV.[6]. Since this variable is ordinal, we use the median instead of the mean across PTAs members. Our expectation is that democracies should sign deeper and more rigid agreements than autocracies. Furthermore, we add a variable that captures if a country has recently undergone a transition from autocracy to democracy ($Democratization$). This variable scores one if *at least* one country has become a democracy over the past ten years using Polity IV as indicator. The expectation is that democratizing countries require more flexibility since they face high levels of uncertainty about future states of the world.

We also add a dummy that scores one if *every* PTA member is a member of the WTO (WTO). WTO members tend to implement trade policies that differ from countries that are not part of this international organization (Mansfield and Reinhardt, 2003). They also have *ad hoc* flexibility provisions upon which they can rely. Finally, we include the number of member countries of a PTA ($No.Members$). On the one hand, $No.Members$ accounts for the broader-deeper trade-off argument in international agreements (Gilligan, 2004). On the other hand, Koremenos et al. (2001) argue that the size of membership is a crucial driver of flexibility. Table 1 summarizes the descriptive statistics of the dependent and independent variables.

Table 1 About Here

Empirical Results

Baseline Analysis

We start by presenting the results for the baseline analysis with ordered probit and tobit estimations. Table 6 shows the results of these baseline models with the findings generally supportive of our hypotheses. Indeed, the coefficient of both *Depth* and *Long-term Flexibility* are positive and statistically significant at the conventional level. Thus, increasing

[6]Results do not change if we use other measures of democracy such as Cheibub et al. (2010).

the depth of a PTA has a positive effect on its long-term flexibility. Similarly, increasing the long-term flexibility of a PTA has a positive impact on its rigidity on flexibility. Although these are simple correlations, paired with the scatter plot showed above, they offer encouraging preliminary support of our theory.

Table 6 About Here

Importantly, control variables have the expected sign when they are statistically significant at the conventional level. Thus, using PTA as unit of analysis does not seem to distort the impact of monadic and dyadic variables on the design of PTAs.

Instrumental Variables

We now move to the core test of our theory. We estimate instrumental variable models, which allow us to endogenize *Depth* and *Long-term Flexibility*. The results showed in Table 3 support our expectations concerning the endogenous determinants of lobbying. First, the coefficient of *Depth* has the expected positive sign in every model when *Long-term Flexibility* and *Short-term Flexibility* are the dependent variables and is statistically significant at the 99 percent confidence level. Second, the coefficient of *Long-term Flexibility* has the expected positive sign when *Rigidity on Flexibility* is the dependent variable and is statistically significant at the conventional level. The findings thus are completely in line with our theory that emphasizes the importance of the endogenous determinants of the design of PTAs.

Table 3 About Here

In Table 4, we summarize the magnitude of the substantive effects of Model 8. Moving *Depth* from its minimum to its maximum value increases by 88 percent the probability of having the highest value of *Long-term Flexibility*, i.e. four. In Table 5, we summarize the magnitude of the substantive effects of Model 10. Moving *Long-term Flexibility* from its minimum to its maximum value decreases by 88 percent the probability of having the lowest value of *Rigidity on Flexibility*, i.e. five. Results of Model 12 show that moving *Depth* from its minimum to its maximum value increases the tariff transition period by 6 years. The substantive effects of our main variables are so large that in explaining flexibility of PTAs a model that neglects the role of *Depth* suffers from serious missing variable problems.

Table 4 and Table 5 About Here

Finally, control variables have the expected sign when they are statistically significant. Importantly, instruments have the expected sign and they are statistically significant at the conventional level in the first stage. This result reinforces our identification strategy. Finally, the results on ρ lead us to reject the hypothesis that the two equations are independent from each other.

Additional Evidence

So far we have argued and convincingly showed that *Depth* is a crucial driver of different types of flexibility, independent of whether we use multivariate regressions or instrumental variables. In this section, to better account for the possibility that depth and flexibility may be interdependent, we simultaneously estimate *Depth* and *Long-term Flexibility* as well as *Depth* and *Short-term Flexibility*. Ideally, we would do so with a simultaneous equation model, which would allow us to estimate the reduced-form equations that result from substituting the expression for each component of the design into the other two components. Unfortunately, in our setting such a model does not converge due to a low number of observations. Thus, we implement an admittedly second-best strategy using seemingly unrelated regression (henceforth SUR). In a SUR model, the two equations have contemporaneous cross-equation error correlation, that is, the error terms in the regression equations are correlated. More formally, we estimate the following model for *Depth* and *Long-term Flexibility*:

$$LongTermFlexibility_i = \alpha_1 + \beta_1 Depth_i + \beta_2 X_i + \beta_3 Z_1 + \epsilon_1. \tag{4}$$

$$Depth_i = \alpha_2 + \beta_4 LongTermFlexibility_{ij} + \beta_5 X_{ij} + \beta_6 Z_2 + \epsilon_2. \tag{5}$$

And similarly for *Depth* and *Short-term Flexibility*:

$$ShortTermFlexibility_i = \alpha_3 + \beta_7 Depth_i + \beta_8 X_i + \beta_9 Z_3 + \epsilon_3. \tag{6}$$

$$Depth_i = \alpha_4 + \beta_{10} LongTermFlexibility_{ij} + \beta_{11} X_{ij} + \beta_{12} Z_2 + \epsilon_4. \tag{7}$$

We have already presented the variables capturing the PTA design, the instruments for *Depth* and *Long-term Flexibility*, and the control variables. In addition, we need to instrument *Short-term Flexibility* in a SUR model. We use the standard deviation of GDP among PTA member countries as a proxy for asymmetric bargain power within a PTA. Weaker states feel that they need long transition periods to address disadvantages

related to relative competitiveness. As agreements are driven by the idea of 'reciprocal commitments' the stronger states will demand similar lengths in transition. Therefore, with increasing market power asymmetry and important differences in the level of competitiveness, we should witness longer transition periods. *Asymmetric Power* is strongly correlated with *Short-term Flexibility*, $\rho = .35$, and weakly correlated with the error term of the *Depth* equation, $\rho = .1$.[7]

Table 2 shows the results of the baseline model. The findings confirm that *Depth* has a positive and statistically significant effect on both *Long-term Flexibility* and *Short-term Flexibility*. The results also show that both *Long-term Flexibility* and *Short-term Flexibility* have a positive and statistically significant impact on the depth of PTAs. Thus, more flexible PTAs tend to be deeper. In other words, in order to be able to afford deep PTAs, member countries need to include flexibility devices that mitigate distributional concerns and spread adjustment costs over a long period of time.

Table 2 About Here

Not only are the effects statistically significant, but their magnitudes are also sizeable. Moving from a PTA with no escape clauses to a PTA with four escape clauses increases *Depth* by more than 9 points. Given that *Depth* ranges between zero and 40, the magnitude of this effect is remarkable. We obtain a similar effect if we look at *Short-term Flexibility*. Specifically, moving from the smallest to the largest value of *Short-term Flexibility* increases *Depth* by 9 points, whereas moving from a standard deviation below the mean to a standard deviation above the mean increases *Depth* by more than 3 points. In sum, flexibility is a crucial driver of the depth of PTAs.

We conclude by noting that the Breusch-Pagan test leads us to reject the hypothesis that the two equations are independent from each other. Thus, a SUR model complements nicely our previous analyses by providing further evidence that member countries design PTAs taking into account different dimensions of treaties. However, we note that a SUR model is not as well-suited as an instrumental variable approach to account for the endogeneity problem.

Robustness Checks

We perform several other tests to check the robustness of our findings. First, we replace the indicator of depth obtained by summing provisions with the indicator of depth using factor analysis (see the discussion above). Second, we use bivariate probit regressions

[7] If we run classic instrumental variable regressions in equations in which *Depth* is the dependent variable, the Cragg-Donald Wald F test of the first stage is 12.97 when *Long-term Flexibility* is instrumented and 28.42 when *Short-term Flexibility* is instrumented. Thus, there is no concern that our instruments are weak.

having transformed the flexibility variables and depth into dummy variables. Third, we implement two-stage estimations "by hand". Specifically, we predict *Depth* in the first stage using a tobit model and we include the predicted values on the right-hand side of the second stage in which the dependent variables are *Long-term Flexibility* and *Short-term Flexibility*.[8] We do the same with *Long-term Flexibility* and *Rigidity on Flexibility*. Finally, we drop from the analysis EU treaties, EU PTAs with third countries, and US PTAs to check if the results are driven by some specific agreements. For all these checks, the results are similar to the ones reported above and are available upon request.

Conclusion

We have argued that international agreements that are deep also need to be flexible. That is, with increasing international obligations, the demand for flexibility goes up. In addition, we have argued that states can offer multiple forms of flexibility mechanisms. In respect to deep trade agreements these include both what we call short-term and long-term flexibility tools. We also expected, however, that faced with exporter demands, states will impose some rigidity on the use of flexibility mechanisms, but leave others unchanged. Put differently, we have posited that while increasing depth translates into a growing demand for flexibility clauses, states also design limits for using long-term flexibility provisions.

Drawing on an original dataset of the design of 587 trade agreements, we have found support for our argument. The depth of agreements indeed has an impact on their flexibility. In addition, the results show that as long-term flexibility increases, we see greater restrictions on the use of the flexibility provisions. These results are robust to using instrumental variables, SUR models and changing the operationalization of depth.

Our findings show that different design features of international agreements are not independent of each other. To explain the flexibility of international agreements, it is not sufficient to look at characteristics of the member states or issue characteristics. Instead, flexibility is to a large extent determined by the depth of the international agreement. The results do not only show that flexibility is endogenous to depth, but they suggest further that states by responding to societal demands, may have some strategic leeway to prioritize certain flexibility mechanisms over others. Here of course the (lengthy) negotiation processes allow parties to an international agreement to come up with innovative solutions to address the concerns of import-competing groups and of exporters. In this sense, further research might more explicitly focus on the degree to which flexibility mechanisms substitute or complement each other and serve different purposes.

[8] We use bootstrapped standard errors in the second stage.

Bibliography

Alvarez, Jose E. 2005. *International Organizations as Law-Makers*. Oxford: Oxford University Press.

Baccini, Leonardo. 2010. "Explaining Formation and Design of EU Trade Agreements: The Role of Transparency and Flexibility." *European Union Politics*, 11(2): 195-217.

Baccini, Leonardo and Andreas Dür. 2012. "The new regionalism and policy interdependence." *British Journal of Political Science*, 42(1): 57-79.

Baier, Scott L. and Jeffrey H. Bergstrand. 2007."Do Free Trade Agreements Actually Increase Members' International Trade?" *Journal of International Economics* 71(1):72–95.

Börzel, Tanja A. and Thomas Risse. 2012. "When Europeanisation Meets Diffusion: Exploring New Territory.", *West European Politics* 35(1): 192–207.

Chase, Kerry A. 2005. *Trading Blocs: States, Firms, and Regions in the World Economy*. Ann Arbor: University of Michigan Press.

Cheibub, Jose A., Jennifer Gandhi and James R. Vreeland. 2010. "Democracy and dictatorship revisited." *Public Choice* 143(1-2): 67–101.

Dür, Andreas. 2010. *Protection for Exporters: Power and Discrimination in Transatlantic Trade Relations, 1930-2010*. Ithaca: Cornell University Press.

Dür, Andreas, Leonardo Baccini, Manfred Elsig and Karolina Milewicz. 2013. "The design of international trade agreements: Introducing a new dataset." Unpublished paper.

Elsig, Manfred and Cédric Dupont. 2012. "European Union meets South Korea: Bureaucratic interests, exporter discrimination and the negotiations of new trade agreements." *Journal of Common Market Studies* 50(3):492-507.

Gilligan, Michael J. 1997. "Lobbying as a private good with intra-industry trade." *International Studies Quarterly* 41(3): 455–474.

Gilligan, Michael J. 2004. "Is There a Broader-Deeper Tradeoff?" *International Organization* 58(3): 459-84.".

Goldstein, Judith, Miles Kahler, Robert O. Keohane and Anne-Marie Slaughter. 2000. "Introduction: Legalization and world politics." *International Organization* 54(3): 385–399.

Goodman, Ryan and Derek Jinks. 2004. "How To Influence States: Socialization and International Human Rights Law." *Duke Law Journal* 54(3):621–704.

Greene, William H. 2011. *Econometric Analysis*, 7th ed. (Harlow: Pearson Education).

Haftel, Yoram Z. 2013. "Commerce and Institutions: Trade, Scope, and the Design of Regional Economic Organizations." *Review of International Organizations*.

Hall, Richard L. and Alan V. Deardorff. 2006. "Lobbying as Legislative Subsidy.",

American Political Science Review 100(1): 69-84.

Helfer Laurence R. 2005. "Exiting Treaties." *Virginia Law Journal* 91(7): 1579–1648.

Helfer, Laurence R. 2013. "Flexibility in International Agreements." In Dunoff Jeffrey and Pollack, Mark (eds.) *Interdisciplinary Perspectives on International Law and International Relations*. Cambridge: Cambridge University Press, 175–196.

Johns, Leslie (2012) "Depth versus Rigidity in the Design of International Trade Agreements." Unpublished paper.

Koremenos, Barbara. 2005. "Contracting around international uncertainty." *American Political Science Review* 99(4): 549–565.

Koremenos Barbara and Allison Nau. 2010. "Exit, No Exit." *Duke Journal of Comparative and International Law* 81(1):81–120.

Koremenos, Barbara, Charles Lipson and Duncan Snidal. 2001. "The rational design of international institutions." *International Organization* 55(4): 761-799.

Kucik, Jeffrey. 2012. "The Domestic Politics of Institutional Design: Producer Preferences over Trade Agreement Rules." *Economics & Politics* 24(2):95-118.

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

Manger, Mark.2009. *Investing in Protection: The Politics of Preferential Trade Agreements Between North and South*. Cambridge: Cambridge University Press.

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(4): 829–862.

Mansfield, Edward D., Helen V. Milner and B. Peter Rosendorff. 2002. "Why democracies cooperate more: Electoral control and international trade agreements." *International Organization* 56(3): 477–513.

Mansfield, Edward D. and Eric Reinhardt. 2008. "International Institutions and the Volatility of International Trade." *International Organization* 62(4): 621–652.

Maskus, Keith E. and Mohan Penubarti. 1995. "How Trade-Related Are Intellectual Property Rights¿' *Journal of International Economics* 39(3): 227–248.

Mearsheimer, John J. 1994. "The False Promise of International Institutions.", *International Security* 19 (3): 5–49.

Milner, Helen V. 1988. *Resisting Protectionism: Global Industries and the Politics of International Trade*. Princeton: Princeton University Press.

Neumayer Eric (2007). "Ratification: Explaining Reservations to International Human Rights Treaties." *Journal of Legal Studies* 36(2): 397–430.

Pelc, Krzysztof J. 2009. "Seeking escape: The use of escape clauses in international trade agreements." *International Studies Quarterly* 53(2): 349–368.

Pelc, Krzysztof J. 2010. "Eluding efficiency: why do we not see more efficient breach at the WTO?" *World Trade Review* 9 (4): 629–42.

Piermartini, Roberta and Michele Budetta. 2009 "A Mapping of Regional Rules on Technical Barriers to Trade." In Estevadeordal, Antoni, Kati Suominen and Robert Teh (eds.), *Bank Regional Rules in the Global Trading System.* Cambridge: Cambridge University Press, 250–315.

Revelle, William. 2012. Package psych, Version 1.2.8.

Roodman, D. 2011 "Estimating fully observed recursive mixed-process models with cmp." *Stata Journal* 11(2): 159-206.

Rosendorff, B. Peter. 2005. "Stability and rigidity: Politics and design of the WTO's dispute settlement procedure." *American Political Science Review* 99(3): 389–400.

Rosendorff, B. Peter and Helen Milner. 2001. "The optimal design of international trade institutions: Uncertainty and escape." *International Organization* 54(4): 829–857.

Shaffer, Gregory C. and Mark A. Pollack 2010. "Hard vs. Soft Law: Alternatives, Complements and Antagonists in International Governance." *Minnesota Law Review* 94(3):706–799.

Stein, Arthur. 1982. "Coordination and Collaboration: Regimes in an Anarchic World.", *International Organization* 36(2): 299–324.

Svolik, Milan. 2006. "Lies, Defection, and the Pattern of International Cooperation." *American Journal of Political Science* 50(4): 909–25.

Train, K. 2009 *Discrete Choice Methods with Simulation*, 2nd ed. Cambridge: Cambridge University Press.

World Bank. 2012. World Development Indicators.

World Trade Organization 2011. *World Trade Report - The WTO and Preferential Trade Agreements: From Co-existence to Coherence.* Geneva: World Trade Organization.

Wright, John R. 1996. *Interest Groups and Congress: Lobbying, Contributions, and Influence.* Boston: Allyn and Bacon.

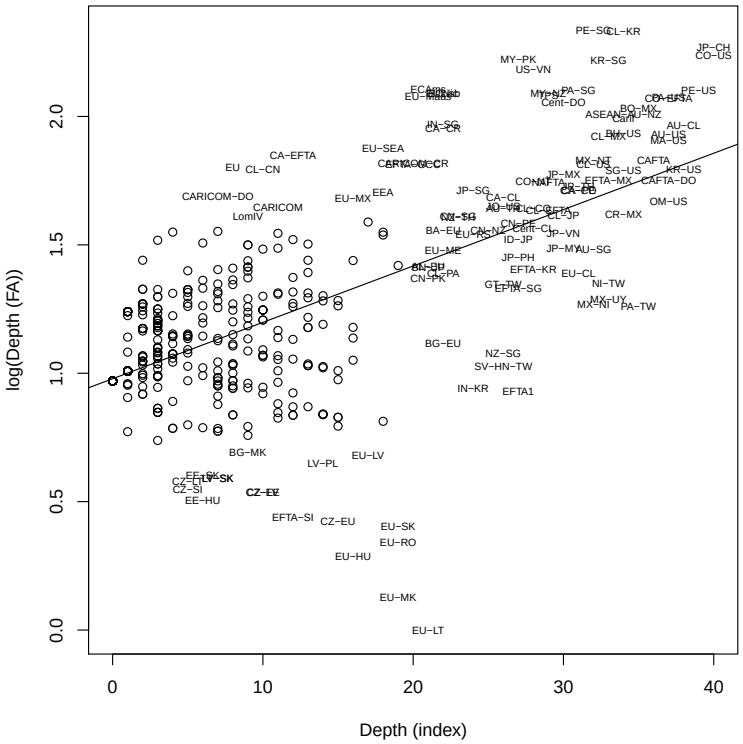

Figure 1: Depth.

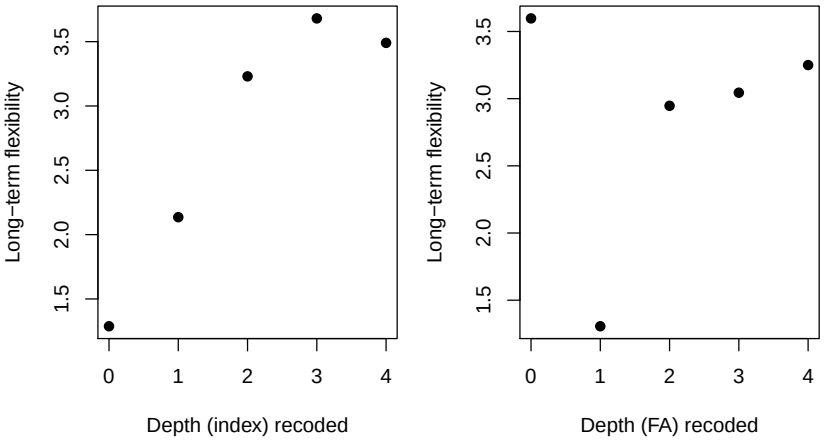

Figure 2: Depth versus long-term flexibility.

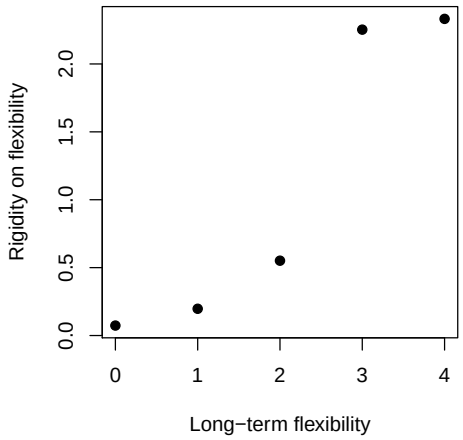

Figure 3: Long-term flexibility versus rigidity on flexibility.

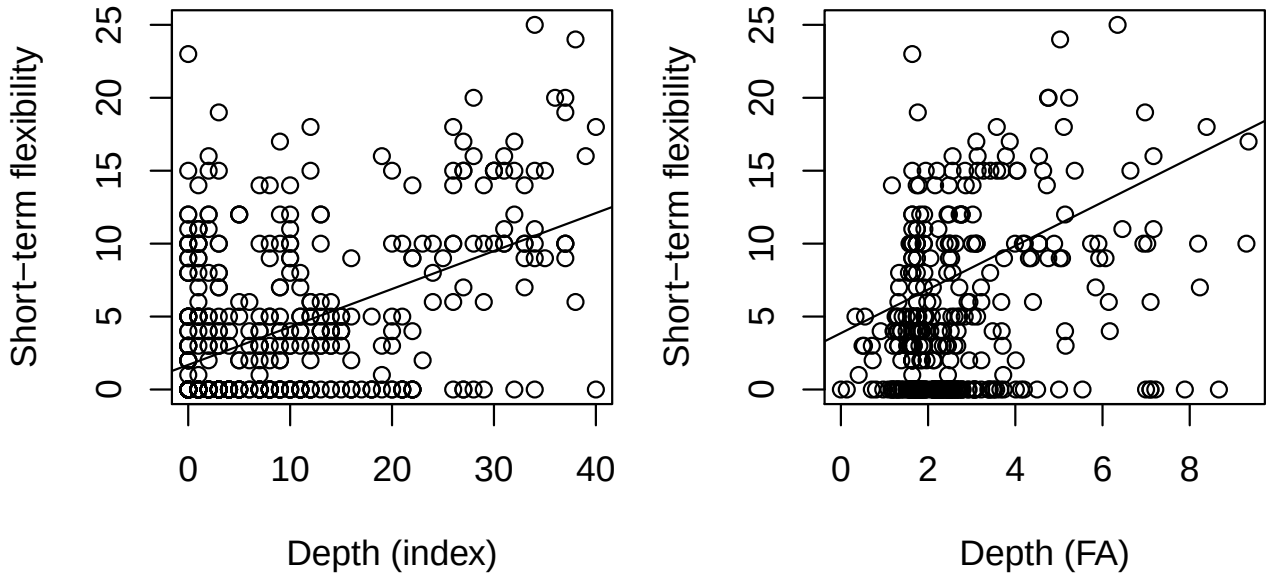

Figure 4: Depth versus short-term flexibility.

Table 1: Descriptive statistics.

Variables	Mean	Std. Dev.	Min	Max
Long-term Flexibility	2.68	1.46	0	4
Rigidity on Flexibility	1.53	1.35	0	5
Short-term Flexibility	3.84	5.21	0	25
Depth	8.13	10.03	0	40
GDPpc	9.10	0.86	5.91	10.83
GDP	23.54	1.83	17.34	27.80
Trade	4.97	2.35	0.06	12.03
Democracy	4.74	5.82	-10	10
Democrat.	0.85	0.36	0	1
WTO	0.50	0.50	0	1
Distance	7.99	.99	4.47	9.87
No. Members	5.65	8.65	2	91
Transparency	0.68	0.30	0	1
WTO Depth	1.31	1.30	-0.75	2.33
EU Depth	4.39	1.10	2.69	6
Asymmetric Power	86.12	214.98	$4.67e^{-15}$	1588.79

Table 2: Baseline models.

Covariates	(1) Long-term Flexibility	(2) Long-term Flexibility	(3) Rigidity on Flexibility	(4) Rigidity on Flexibility	(5) Short-term Flexibility	(6) Short-term Flexibility
Depth	0.05***	0.05***			0.36***	0.34***
	(0.01)	(0.01)			(0.04)	(0.04)
LongTermFlexibility			0.76***	0.80***		
			(0.05)	(0.05)		
GDP	-0.06	-0.07	0.04	0.06	-0.31	-0.48
	(0.04)	(0.05)	(0.05)	(0.05)	(0.36)	(0.37)
GDPpc	0.36***	0.27***	0.15***	0.13	-0.75	-0.71
	(0.08)	(0.08)	(0.09)	(0.09)	(0.61)	(0.66)
Trade	-0.03	-0.03	0.14***	0.11***	0.66***	0.64**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.26)	(0.26)
No.Members		-0.01		-0.01		0.08*
		(0.01)		(0.01)		(0.05)
Regime		0.04***		-0.01		-0.06
		(0.01)		(0.01)		(0.08)
WTO		0.05		0.61***		2.77***
		(0.10)		(0.11)		(0.90)
Democratization		0.39***		-0.03		0.70
		(0.01)		(0.13)		(1.11)
Cut 1	0.75	0.15	4.53***	4.95***		
	(0.83)	(0.93)	(0.87)	(0.97)		
Cut 2	1.29	0.71	5.29***	5.75***		
	(0.83)	(0.92)	(0.87)	(0.98)		
Cut 3	1.79**	1.22	6.14***	6.64**		
	(0.83)	(0.93)	(0.88)	(0.99)		
Cut 4	2.29***	1.75	7.46***	8.00***		
	(0.83)	(0.93)	(0.89)	(1.00)		
Cut 5			8.49***	9.07***		
			(0.91)	(1.01)		
Constant					8.42	9.97
					(7.12)	(7.73)
Observations	565	561	548	545	497	494
Pseudo R-squared	0.08	0.10	0.26	0.28	0.05	0.06

Models 1, 2, 3, and 4 are ordered probit. Models 5 and 6 are tobit.
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Instrumental variable models.

Covariates	(7) Long-term Flexibility	(8) Long-term Flexibility	(9) Rigidity on Flexibility	(10) Rigidity on Flexibility	(11) Short-term Flexibility	(12) Short-term Flexibility
$\widehat{Depth}$	0.10***	0.10***			0.16**	0.14*
	(0.01)	(0.01)			(0.08)	(0.08)
$LongTerm\widehat{Flexibility}$			1.02***	1.07***		
			(0.05)	(0.05)		
GDP	-0.06	-0.07	0.06	0.08*	-0.29	-0.49
	(0.04)	(0.04)	(0.04)	(0.05)	(0.37)	(0.38)
$GDPpc$	0.09	0.06	-0.06	-0.04	0.03	-0.01
	(0.08)	(0.09)	(0.09)	(0.09)	(0.68)	(0.71)
$Trade$	-0.11***	-0.09	0.11***	0.08**	0.99***	0.93***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.29)	(0.29)
$No.Members$		-0.01		-0.001		0.08*
		(0.01)		(0.01)		(0.05)
$Regime$		0.03***		-0.03***		-0.04
		(0.01)		(0.01)		(0.08)
WTO		-0.20*		0.45***		3.49***
		(0.11)		(0.12)		(0.95)
$Democratization$		0.34**		-0.16		0.98
		(0.13)		(0.15)		(1.10)
Cut 1	-1.43	-1.48	3.53***	4.19***		
	(0.95)	(0.93)	(0.88)	(0.97)		
Cut 2	-0.97	-1.00	4.20***	4.90***		
	(0.86)	(0.93)	(0.90)	(0.98)		
Cut 3	-0.54	-0.55	4.97***	5.70***		
	(0.87)	(0.94)	(0.91)	(1.00)		
Cut 4	-0.10	-0.09	6.16***	6.93***		
	(0.88)	(0.94)	(0.94)	(1.03)		
Cut 5			8.49***	7.90***		
			(0.97)	(1.05)		
Constant					0.79	3.37
					(7.64)	(8.15)
$WTODepth$	1.75***	1.47***			3.22***	2.95***
	(0.41)	(0.41)			(0.63)	(0.63)
$EUDepth$	1.74***	2.01***			2.34***	2.70***
	(0.50)	(0.52)			(0.78)	(0.78)
$Transparency$			-1.75**	-1.68**		
			(0.70)	(0.73)		
$Transparency^2$			2.19***	1.75***		
			(0.63)	(0.65)		
Observations	565	561	548	545	497	494
Rho	-0.58***	-0.55***	-0.55***	-0.57***	0.52***	0.51***

Models 1, 2, 3, and 4 are ordered probit (instrumented). Models 5 and 6 are tobit (instrumented)
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Predictions from Model 8: Depth on long-term flexibility.

Value	Min → Max	C.I.
0	-0.65	[-0.81, -0.48]
1	-0.16	[-0.22, -0.11]
2	-0.07	[-0.12, 0.04]
3	0.01	[0.005, 0.03]
4	0.88	[0.78, 0.95]

Table 5: Predictions from Model 10: Long-term Flexibility on Rigidity on Flexibility.

Value	Min → Max	C.I.
0	-0.88	[-0.92, -0.83]
1	0.09	[0.05, 0.13]
2	0.31	[0.26, 0.26]
3	0.40	[0.35, 0.46]
4	0.08	[0.05, 0.11]
5	0.01	[0.002, 0.02]

Table 6: SUR models.

Covariates	(1) Long-term Flexibility	(2) Depth	(3) Short-term Flexibility	(4) Depth
Depth	0.08***		0.39***	
	(0.01)		(0.03)	
LongTermFlexibility		2.35***		1.03***
		(0.24)		(0.06)
GDP	-0.05	0.44	-0.04	0.35
	(0.05)	(0.29)	(0.22)	(0.32)
GDPpc	0.38***	1.52***	-1.61***	3.58***
	(0.09)	(0.50)	(0.35)	(0.52)
Trade	-0.08**	0.85	0.07	0.15
	(0.04)	(0.20)	(0.15)	(0.23)
Transparency	-1.43*			
	(0.80)			
$Transparency^2$	1.69**			
	(0.74)			
WTODepth		1.16**		1.49***
		(0.45)		(0.46)
EUDepth		1.39**		0.98*
		(0.58)		(058)
AsymmetricPower			0.002**	
			(0.001)	
Constant	0.19	-34.33***	15.90***	-48.49***
	(1.01)	(5.64)	(4.23)	(6.12)
Observations	518		455	
Pseudo R-squared	0.21	0.43	0.20	0.51
Breusch-Pagan test of independence	36.73***		65.76**	

SUR models. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Appendix

Operationalizing depth

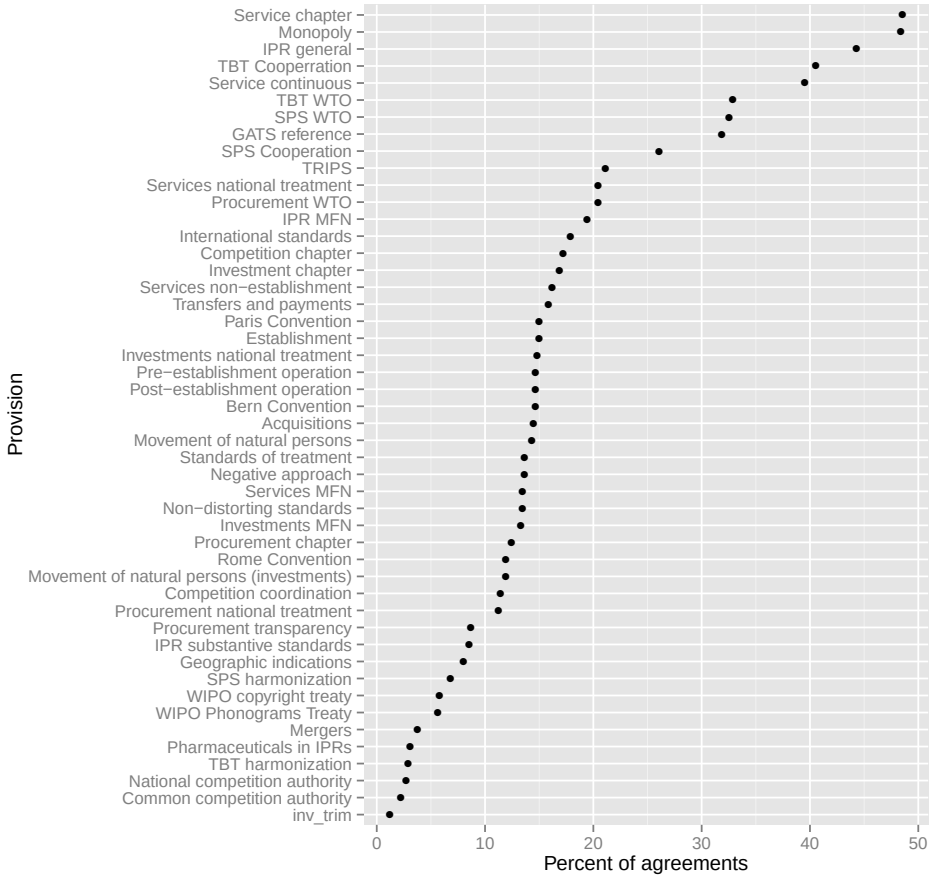

Figure 5: The items included in depth.

Operationalizing long-term flexibility

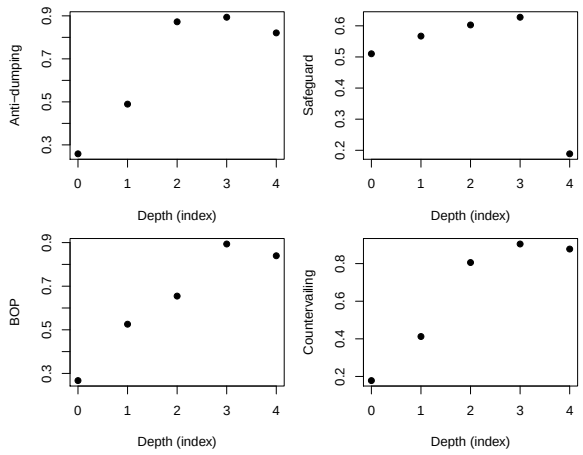

Figure 6: The items included in long-term flexibility.

Operationalizing rigidity on flexibility

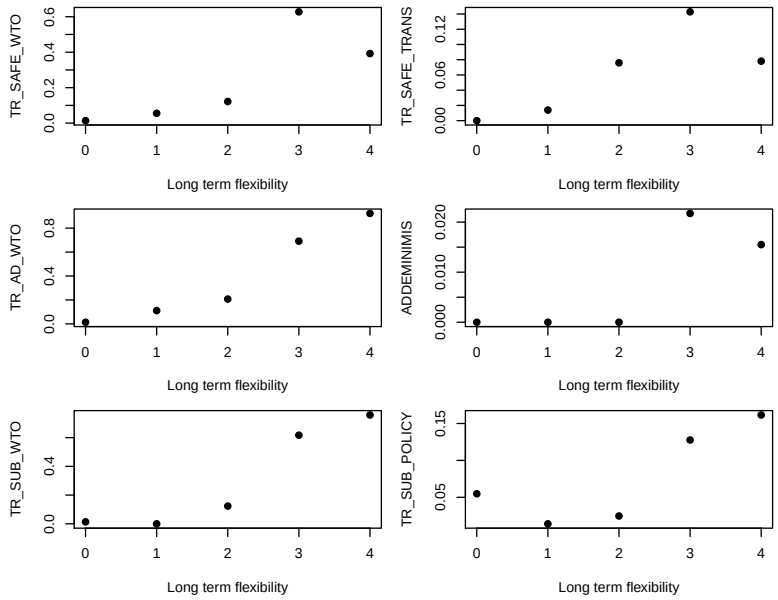

Figure 7: The items included in rigidty on flexibility.

The instruments

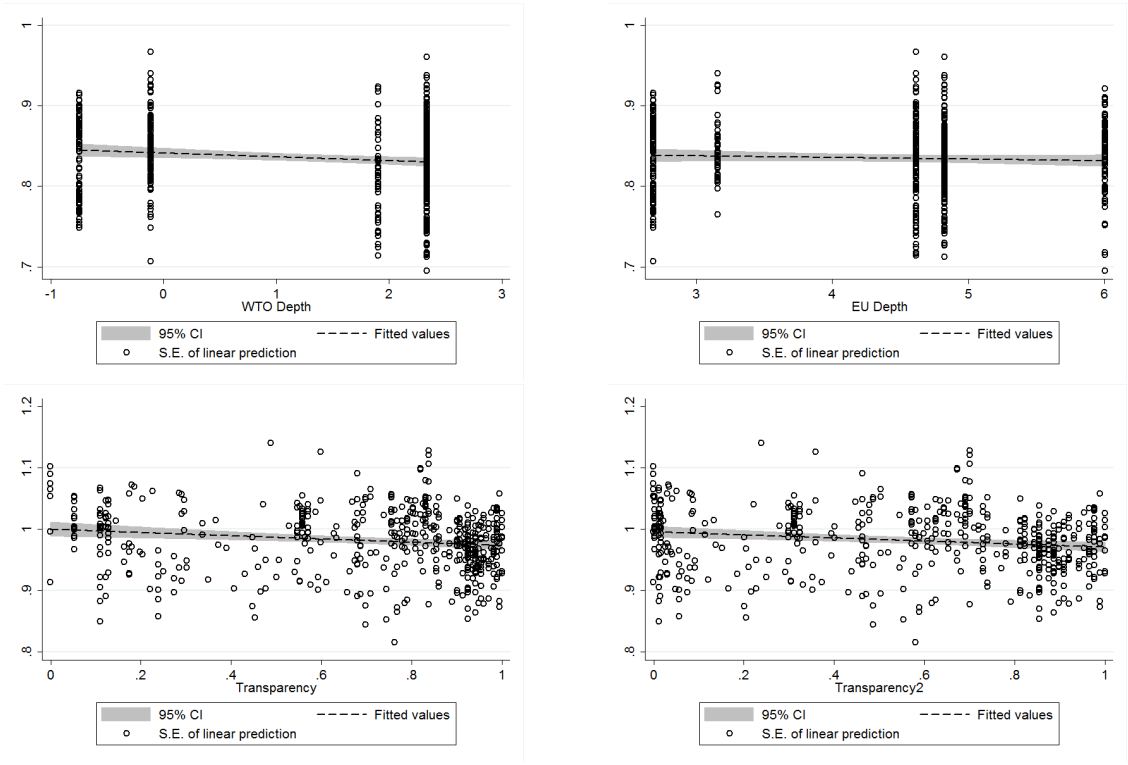

Figure 8: Correlations between instruments and residuals