Leadership Turnover and the Implementation of International Economic Agreements

Abstract

Under what conditions do states implement their international agreements? While recent work focuses on the formation and design of agreements, we know comparatively little about the factors that shape implementation. Yet, this issue is crucial for understanding how (or whether) international agreements promote cooperation. In this paper, we explore how changes in domestic political conditions shape implementation. We argue that international agreements are implemented less fully in the wake of government turnover. New leaders are less willing to abide by a previous government's commitments, particularly when the ideological leanings of the chief executive are different from the previous regime. Our argument therefore casts doubt on the traditional view that leaders enter into international agreements to lock their preferred policies in place for future governments. Instead, changes in domestic leadership have perceptible effects on international cooperation. We test this argument using data on around 300 preferential trade agreements (PTAs) since 1970. Ideological turnover is shown to have a negative and statistically significant relationship with PTA implementation. Our results are robust to a variety of model specification, including attempts to correct for the non-random selection process by which states sign PTAs in the first place. These findings have important implications for our understanding of international cooperation: if states do not adhere to the commitments that they negotiate, the bargaining stage of agreements may not have lasting consequences for economic cooperation.

1 Introduction

Nearly every country in the world has signed an economic cooperation agreement with another state. But how do those agreements fare once the executive who signed the agreement leaves office? Existing work argues that international agreements tie politicians' hands — and crucially, that these constraints extend beyond an agreement's original signatories.[1] According to many theories of cooperation, a core virtue of formal agreements is that they "lock in" agreed-on policies, making it difficult for any future leaders to violate the contract.[2] If this view is correct, then once an agreement is signed, the durability of international cooperation should not be linked to individual leaders. Instead, agreements ought to shape policy even when there are changes in the domestic political climates of member states. But for the most part, this claim has not been systematically tested.

This paper argues that changes in domestic leadership fundamentally shape the implementation of international agreements. We posit that the implementation of economic agreements — the focus of this analysis — is less likely when political turnover results in a shift in a member government's policy preferences, which we measure as changes in the ideological leanings of the chief executive. When a new leader has a different ideological predisposition than the previous executive, we expect that leader to distance the government from the *status quo*. In practice, the terms of economic agreements reflect a distributional bargain that may help certain domestic constituents at the expense of others. This bargain is likely to be viewed less favorably as a new leader, with a different set of preferences, enters office. The observable implication is that we a change in leadership results in less agreement implementation, when that turnover is accompanied by a change in ideology.

Understanding whether new governments fully implement the commitments made by their predecessors deepens our understanding of how institutions matter, particularly in light of concerns about implementation and compliance with IOs that have been raised in the literature.[3] More precisely, examining the effects of agreement membership on future governments' behavior helps us understand agreements' legacies past their initial date of entry into force. Although many

[1] Keohane 1984; North and Weingast 1989; North 1990.
[2] Martin 2000; Moravcsik 2000; Simmons 2000, 2001.
[3] Barnett and Finnemore 1999; Mitchell and Hensel 2007; Haftel 2007.

studies argue that international agreements produce durable commitments, few have investigated this claim systematically. The nature of those commitments is also assumed to be somewhat strict: many researchers assume that simply staying in an organization or not disbanding an agreement is tantamount to international cooperation. However, as the literature on compliance has noted,[4] countries need not disband an agreement altogether in order for cooperation to flag. Although formal exit from international organizations is infrequent,[5] new leaders may simply neglect or fail to act on the previous governments international commitments if they are not a top priority. Thus, even though an agreement may formally pervade across changes in government, it may not be implemented in full.

Although this argument could apply to many types of international agreements, here we focus on preferential trading agreements. PTAs continue continue to proliferate. According to the World Trade Organization (WTO) the current number of PTAs active in the international system is around 583, with 377 in force. These agreements cover the majority of world trade and include countries at all levels of development, and in all regions of the world. As they cover economic activity, their implementation – which we define as the removal of political barriers to trade[6] — is relatively easy to measure, unlike other types of agreements that might have aims that are more difficult to define.

We show that even in the presence of an international economic agreement, new leaders tend to remove the political barriers to trade *less* fully when their ideological leanings diverge from the previous executive. That is to say, they implement agreements to a lesser extent. Across our tests, ideological turnover is associated with a significant persistence of political barriers to trade on average — meaning that protectionist practices tend to endure in the presence of leadership turnover. Our findings are robust to a variety of model specifications, including controls for the various factors that might confound our inferences, as well as the potential for bias resulting from

[4]See, for example, Underdal 1998, who argues that compliance is not binary; states may comply with some parts of an agreement but not others, or meet some requirements only partially. See also Barnett and Solingen 2007; Axline 1994; Allee and Scalera 2014; Colgan 2014.

[5]Vabulas 2014.

[6]Specifically, we operationalize implementation as the residuals from a gravity model of trade between PTA partners, which represent the degree of trade between countries that either meets, exceeds, or falls below the level of trade that would be anticipated based on those countries' geographic and economic fundamentals. More details on this measure are provided in the section on operationalization.

non-random selection into leadership turnover. We take into account the processes that might be driving both political turnover and agreement implementation simultaneously, and find that our results endure.

These findings are consistent with studies that express skepticism about compliance with international agreements.[7] Some scholars have claimed that states merely sign agreements that reflect behavior that they might have engaged in, even in the agreement's absence.[8] Others have pointed out that many agreements, once signed, languish at the ratification stage within signatory countries.[9] In keeping with these themes, this paper shows that international agreements may not constrain the behavior of future politicians. Once electoral turnover is taken into account, the dynamics of international economic cooperation shift toward nonimplementation, as others have argued not just for trade[10] and for other forms of international activity, such as sanctions[11] and sovereign debt repayment.[12] This is a less dramatic form of defection from international agreements than outright abnegation, but it offers a more realistic picture of the powers of international agreements to shape member-state behavior over time.

This paper proceeds as follows. The next section reviews the extant literature on international commitments and leadership turnover, particularly as it applies to economic agreements. We lay out our primary hypotheses as well as rival explanations. Section Three describes the operationalization of the key concepts, including a new measure of agreement implementation: trade between dyads of states that are at or different from what might be predicted at a baseline level given a set of countries' geographic and economic fundamentals. This measure captures the *political* barriers to trade that might exist even if an international agreement is present. Section Four examines the empirical relationship between leadership turnover and agreement implementation, for 164 countries from 1968 to 2004. We find that leadership turnover in one of the countries in a preferential trade agreement is associated with lower levels of implementation of that agreement; this relationship is robust to a variety of specifications as well as to methods that control for potential endogeneity in

[7] Underdal 1998; Tallberg 2002.
[8] Downs, Rocke and Barsoom 1996; Von Stein 2005.
[9] Haftel and Thompson 2013; Chapman and Chaudoin 2013; Kelley and Pevehouse 2014.
[10] McGillivray and Smith 2004.
[11] McGillivray and Stam 2005.
[12] Dhillon and Sjostrom 2009.

the relationship. The final section concludes.

2 International Agreements and Leadership Turnover

A central claim of the literature on international agreements is that they help lock in patterns of behavior for future generations of politicians.[13] By committing to an international treaty, the argument goes, politicians must uphold the commitments delineated in international agreements and are thus constrained in their behavior. An agreement that effectively ensures cooperation at the international level ought to tie the hands of not just present, but also *future* leaders. Forming commitments in the international arena insulates leaders from domestic pressures to backslide or renege.[14]

Researchers frequently apply these theories to the area of economic cooperation. Many argue that international trade agreements resolve the cooperation problems inherent in mutual liberalization, preventing domestic interest groups from lobbying politicians to opt for protection.[15] However, these assumptions are rarely tested directly, despite appearing frequently in the background of many arguments. For example, studies of PTAs often take the rigidity of an agreement as evidence of its constraining behavior;[16] the more rigid an agreement, the more credible of a commitment it will serve to future governments.[17] The assumption of agreements as credible commitments underlies many of the central findings of the cooperation-enhancing effects of PTAs, including reduced volatility [18] and increased investment [19]. Some have also argued that the constraining effects of agreements prevent politicians from losing office.[20]

However, other research points out that many economic agreements have a somewhat poor record of implementation over the course of their existence, or at least a variable one.[21] Haftel 2012, for example, finds a significant "implementation gap" among the economic organizations

[13] Keohane 1984; North and Weingast 1989; North 1990; Martin 2000; Moravcsik 2000; Simmons 2000, 2001.
[14] Mansfield and Pevehouse 2006.
[15] Grossman and Helpman 1995.
[16] Abbott and Snidal 2000; Koremenos, Lipson and Snidal 2001.
[17] Hicks and Kim 2012; Baccini and Urpelainen 2014; Johns 2012.
[18] Mansfield and Reinhardt 2008.
[19] Buthe and Milner 2008.
[20] Mansfield and Milner 2012; Hollyer and Rosendorff 2012.
[21] Eicher and Henn 2008.

that he surveys. Peinhardt and Allee 2012 find that PTAs do not have the investment effects that one would presume from the provisions in those agreements. This broader implementation gap has been the subject of speculation in the studies of international organizations and compliance [22] more generally.[23]

We argue that domestic leadership turnover — particularly when it results in a turnover of the incumbent's party — is one factor that contributes to low levels of implementation in PTAs. Many have noted that leadership turnover impacts the international sphere in many different arenas, including on UN voting[24] and the resolution of WTO disputes,[25] as well as in international conflict.[26]

This pattern extends to the area of international economic agreements as well. Once the leader who negotiates a PTA leaves office, those agreements may be actively set aside, or passively ignored.[27] Any trade agreement represents a bargain with provisions to offset economic losses endured by a particular group, and politicians negotiate to insulate their supporters and constituents from those losses.[28] Thus, a distributive bargain that worked for the signing regime may not be relevant to the constituents of a subsequent leader. Thus, there are many potential political barriers to implementation that might magnify with leadership turnover. Although executives cannot control the extent to which firms actually utilize agreements, they can exercise authority over the broader regulatory environment in which firms operate, making that environment more or less conducive to free trade.

We argue that this is especially true if a leader's ideology diverges from that of the preceding executive. Leaders' ideologies fundamentally shape their policy preferences.[29] When a new leader comes to power after having toppled an incumbent of a different ideology, he or she will not be invested in the policies of the previous ruling party and is likely as well to have new constituents to appease. Rather than take the drastic step of actually disbanding or exiting an agreement —

[22] Simmons 2000; Walter 2008.

[23] On noncompliance in the EU, see Boerzel 2001, 2005; Jensen 2007, but for an opposing view, see Levitz and Pop-Eleches 2010. On noncompliance in Asia, see Ravenhill 2008; Walter 2008; on Africa, see Bourename 2002.

[24] Dreher and Jensen 2013; Smith 2014.

[25] Bobick and Smith 2013.

[26] McGillivray and Smith 2003; Wolford 2007; Saunders 2011; Wolford 2012; Horowitz and Stam 2014.

[27] Smith 2009.

[28] Grossman and Helpman 1995.

[29] Budge and Hofferbert 1990; Schmidt 1996; Bergman, Muller, Strom and Blomgren 2003.

an occurrence that is relatively rare [30] — executives may cave in to protectionist pressures in the form of nontariff barriers such as subsidies, quotas, or laws and regulations that would impede the free flow of goods.[31]

In particular, existing work shows that ideology is an important predictor of whether a leader supports free trade or protection. [32] Parties on the right are associated with preferences for *laissez faire* economics, limiting the role of government in the marketplace. By extension, conservatives often hold a broad commitment to free trade. Conversely, politicians on the left often support policies designed to correct for distributional asymmetries, emphasizing the role governments must play in redistribution. Left-wing parties tend to favor more protectionist policies as a result.

However, we argue that ideological turnover itself, regardless of whether that shift is more to the right or the left, affects implementation. New leaders, if they represent a change from the *status quo*, will seek to distance their policies from those of their predecessor. Leadership change occurs non-randomly; it signals a shift in constituents' underlying preferences. As a result, a new leader has incentives to enact reforms that distance the current government from the previous one. For example, in 2004 Mikheil Saakashvili took the office of the president of the Republic of Georgia. He replaced Eduard Shevardnadze, a former Soviet minister who had maintained close ties to Russia even after the breakup of the Soviet Union. Sheverdnadze previously signed a free-trade agreement in Russia in 1994. However, even though Saakashvili had a far more pro-market ideology than did his predecessor — he liberalized nearly every sector and began an aggressive campaign for FDI — trade with Russia flagged, partially because Saakashvili encouraged firms to establish economic ties to the west rather than the east. Thus, even though Saakashvili had a more right-wing ideology than the former president, implementation of the FTA with Russia stalled because of his attempts to distance himself from the policies of the former regime.

We phrase this proposition as a testable hypothesis:

- $\mathbf{H}_1$ Turnover that results in a shift in the ideology in a member country's executive is nega-

[30] Vabulas 2014.

[31] For example, despite the presence of a PTA with Chile since 1999, the latter brought a WTO dispute that the Mexican government unfairly categorized Chilean matches as hazardous. Such technical nuances in regulation are a widespread form of nontariff barrier.

[32] Nollen and Quinn 1994; Milner and Tingley 2011.

tively associated with the implementation of preferential trade agreements.

There are many mechanisms that might serve as the foundation for this proposed association. For one, delays in ratification can be one means through which agreements fail to launch even after the bargaining process is over.[33] To illustrate, in his second presidential term, the United States Trade Representative under George W. Bush negotiated PTAs with Colombia, South Korea, and Panama. Once Barack Obama took office in 2008, representing a significant shift in the ideological leanings of the executive office, those agreements remained mired in Congress for several years before they were eventually ratified, even though the legislatures in the partner countries had already pushed the agreements through.[34]

Crucially, even if agreements have already made it through the ratification stage, political turnover may mean that new governments simply ignore the agreements that are already in place. For example, new leaders with varying ideologies have caused frequent trouble for Mercosur, a proposed customs union among Brazil, Argentina, Paraguay and Uruguay. First conceived in 1985 by President Jose Sarney of Brazil, the bloc first hit trouble when the next Brazilian president, Fernando Collor de Mello, became embroiled in charges of corruption that later lead to his impeachment. Although de Mello was a committed free trader, his competing domestic troubles stalled the implementation of Mercosur, although the other countries in the agreement continued to push the agenda. The Uruguayan president at the time was quoted as saying, "The integration process is happening between nations, not between political regimes or administrations. This process will continue when I am no longer president and will continue when governments change."[35]

Trade subsequently picked up among the countries after de Mello stepped down, but in 2001 the Argentine financial crisis led to the election or appointment of several different presidents in quick succession, all of whom claimed to reject the neoliberal policies of the previous administrations. During that time period, trade within the bloc contracted and never picked back up, particularly after the election in Argentina of the Kirchner family, who pursued statist policies. In Brazil, the

[33] Haftel and Thompson 2013.

[34] While the agreements did eventually pass, our point is that the process was protracted precisely because of Obama's previous campaign pledges to the contrary. In addition, Obama was heavily criticized for what appeared to be a reversal of his stance.

[35] "Brazil crisis no block to South American common market," United Press International, August 27, 1992

election of Luiz Lula da Silva brought with it a trade policy that looked more toward opportunities outside, rather than within, the region. Mercosur was not disbanded, but Lula focused more on exercising leadership at WTO negotiation rounds and forming new PTAs than he did on dealing with implementation difficulties in Mercosur [36].[37] Thus, even though the agreement itself remained unchanged, its levels of implementation varied with the varying preferences of leaders.

Another way that turnover can result in stalled agreement implementation is if the agreement becomes subject to debate and opposition in the election. For example, a PTA between the US and several Central American countries as well as Costa Rica (known as CAFTA-DR) was negotiated in 2003 and signed by Costa Rica in 2004.[38] The agreement, although it had been been introduced for ratification into the Costa Rican legislature in 2005, became a controversial issue in the 2006 presidential campaign. Subsequent elections resulted in a narrow victory for Oscar Arias of the left-leaning National Liberation Party (PLN), which actually had supported the agreement — but election was sufficiently close that the opposition party pushed for a referendum in 2007, which only barely went through in favor of the agreement.[39]

In sum, many implicitly assume that international agreements bind leaders into certain patterns of behavior. Yet, it remains an empirical question whether leadership turnover in one of the signatory countries has an effect on whether that agreement is actually implemented. Additionally, there is reason to believe that international agreements may fall by the wayside if they do not match the preferences or priorities of newly elected leaders. Rather than exit from the agreement altogether, however, the agreement would simply not be implemented.

2.1 Alternative and Competing Explanations

Our central hypothesis cuts against much of the conventional wisdom about the function of international agreements, which claims that agreements are put in place internationally to solve the

[36] Mahrenbach 2013.

[37] The ideological shift of Mercosur became even more pronounced when Hugo Chavez's Venezuela joined the bloc in 2012. This even went so far as to prompt many to view Mercosur as protectionist; in one article, the *Financial Times* quoted an analyst as saying that member states' "growing disenchantment" with the bloc would mean that members would try "to find ways to bypass Mercosur so they can have different trade agreements. ... For each country, Mercosur will just become less relevant. These agreements never die, they always linger." *Financial Times*, "Mercosur Views Pacific Alliance with Unease," 1 April 2014.

[38] Urbatsch 2013.

[39] Hicks, Milner and Tingley 2014.

time inconsistency problems faced by politicians. The literature on rational design of agreements [40] argues that deep agreements are more likely to be associated with higher levels of cooperation, since countries would not have signed on to an agreement and bargained to produce a deal that asked for levels of cooperation that exceeded their capabilities. If an agreement was not going to be implemented, according to these theories, we would not even observe its existence. Formal agreements specify rules that delineate the boundaries of legally permissable behavior. If international law has any independent effect on states, then these rules constrain the policy options available to member governments. Thus, any finding that agreements remained unimplemented, and that future leaders were unconstrained by international treaties, would stand in opposition to these claims.

Similarly, recent work on the design of international agreements in PTAs[41] takes the presence of a deep agreement as a sign of a credible commitment to a certain trajectory of behavior. During the bargaining process, according to this line of thinking, states build their preferences for cooperation into the PTA itself. Thus, a PTA that asked for high levels of commitment is often assumed to offer a high degree of restraint on member state behavior. However, these arguments do not allow for the possibility that more deep commitments may actually be more likely to be shrugged off by incumbents. That is, if a PTA asks for a high level of commitment, new leaders might be more likely to ignore all or part of the agreement if it does not match their preferences or priorities. Our empirical analysis will investigate whether the design of agreements, particularly as it pertains to the depth of an agreement, makes implementation more or less likely.

Our theory also stands somewhat in contrast to agruments that the signing of preferential trade agreements helps leaders stay in office, either because voters favor the benefits brought by free trade[42] or by neutralizing domestic opposition groups.[43] The implication of this argument in relation to ours is that, if voters favored the economic gains brought on by free trade, they would reward politicians for signing PTAs by casting votes in their direction. These arguments would predict that voters and interest groups would actually prefer for PTAs to be implemented,

[40] Koremenos, Lipson and Snidal 2001.
[41] Rosendorff and Milner 2005; Hofmann and Kim 2014; Dür, Baccini and Elsig 2014.
[42] Mansfield and Milner 2012; Hollyer and Rosendorff 2012.
[43] Brown and Urpelainen 2014.

and would subsequently strike down leaders who failed to do so. This points to the potential for endogeneity in the relationship that we describe. It could be the case that politicians who do not implement PTAs perform poorly in terms of their obligations to voters more generally. Leadership turnover is not randomly distributed; citizens may vote out leaders who fail to fulfill their commitments on a variety of dimensions, or they may experience international pressure to do so.[44] Thus, the relationship between leader turnover and implementation could run the opposite way: poor implementation could be driven by the same underlying factors that simultaneously promote leadership turnover. We investigate this possibility in the empirical portion of the paper.[45]

Other factors might condition the relationship we described. For example, executives have differing impacts across political systems. They have more leverage in presidential systems than in parliamentary systems,[46] and significantly more leverage in autocratic regimes.[47] The types of coalitions necessary to enact change could also mitigate the effect of leaders; their ability to affect policy will differ in majoritarian systems than in proportional representation.[48] Thus, we might expect political turnover to have more of an impact on implementation conditional on regime type and political system.

There are also many processes separate from political turnover that might lead to poor implementation of agreements. There might be factors that simultaneously drive political turnover as well as poor implementation. Economic shocks or contractions of the economy, for example, would be likely to both create voter dissatisfaction that might lead to incumbents being thrown out of office and would simultaneously result in economic agreements not being implemented. If this were the case, it would be necessary to first model the likelihood of political turnover in a given country before examining turnover's effect on the implementation of economic agreements. We will control for all of these possibilities in the empirical analysis.

[44] Marinov 2005; Licht 2010.

[45] There is also some possibility that implementation is influenced by leaders' nonrandom selection into agreements, but those arguments would support agreements being implemented, not unimplemented. Concerns of selection bias center on the idea that leaders only sign agreements that reflect the behavior that they would have engaged in even in the absence of that agreement (Vreeland, 2003; Von Stein, 2005; Mitchell and Hensel, 2007). But this would predict high, not low, compliance with agreements; the argument would not predict that leaders would sign agreements that would not subsequently be implemented, since they would deliberately choose agreements that were in accordance with their intentions. This is contrary to our empirical claim.

[46] Lijphart 1999.

[47] Cheibub, Gandhi and Vreeland 2010.

[48] Austen-Smith and Banks 1988; Torsten Persson 2005.

3 Operationalization and Data Analysis

To test our hypotheses, we gathered data on the leaders of countries in PTAs and on the trade relations among members from the years 1970-2010. The data include roughly 300 agreements formed over this period, involving 800 leaders across 150 countries.[49] The sampled countries exhibit wide variation regarding regime type and level of development.[50] As a result, there is significant diversity among the countries in the sample.

Since we are interested in the conditions under which states honor their PTA commitments, our unit of analysis is country-PTA-year. This choice assumes that when new leaders come to office, they will not necessarily renege on all of their international obligations indiscriminately. Instead, there may be certain features of particular agreements that might make leaders more or less likely to adhere to those commitments. Thus, there is one row for each country i, in PTA p, in year t.[51] Each country-PTA panel includes years starting from the year the agreement enters into force until the disagreement dissolves, if ever.

Dependent Variable. Testing our hypotheses requires a measure of PTA implementation. Although many PTAs have multiple provisions and broad scope that can cover non-economic arenas, PTAs are at their core formed to boost trade.[52] Research in economics supports this assumption, finding that PTAs increase intra-agreement trade by as much as 50 percent.[53][54]

Relying simply on trade levels, however, does not provide a yardstick for measuring implementation at the political level. There are, of course, the problems of identifying any agreement's independent effect on trade, given the possibility of selection bias into agreements that reflect coun-

[49] The average tenure of these leaders is 5.8 years (SD: 7.11)

[50] The average Polity score of the sample is 4.22 (SD: 6.44) and the average level of development is $12,292 per capita (SD: $9,512).

[51] Our unit of analysis means that there are multiple observations per each country-year when country i is a member of multiple PTAs.

[52] Foster, Poeschl and Stehrer 2011.

[53] Baier, Bergstrand and Feng 2013.

[54] That said, the evidence on trade promotion is mixed, and not all agreements are equally beneficial. Moreover, there are agreements that appear to be formed primarily for "political" reasons — e.g. the PTAs between the United States and Bahrain, Oman, and Jordan. Given the lack of complementarities between these markets, it is more likely that the US formed such PTAs to advance its foreign policy agenda. However, even in these cases, the promotion of trade would strengthen the diplomatic ties between members, and is therefore a desirable outcome. Our point is that this variation is shaped fundamentally by variation in members' levels of implementation.

tries' intended behavior even absent the agreement itself. But more importantly, we would need to identify the components of economic exchange that are in the realm of political control. Trade flows themselves reflect the *utilization* of an agreement by firms. But we need a measure of the persistence of political impediments to trade that would be within the authority of politicians to revoke.

Thus, we generate our dependent variable by taking the residuals from a standard gravity model. The gravity model provides us with a prediction of the volume of trade a particular pair of states ought to conduct, given their economic and geographic fundamentals as well as other ties.[55] We cannot know whether a given amount of trade signals anything about implementation without first knowing how much a set of countries *ought* to trade, given their economic and geographic fundamentals and absent political barriers to trade. Subtracting that expectation from the observed value — the residuals of the model — tells us whether a given pair of states is trading more or less than it should. We run of a model of the following form:

$$\begin{aligned} Imports_{i,j,y} \;=\; & \beta_0 \;+\; \beta_1 * lnGDP_{i,t} \;+\; \beta_2 * lnGDP_{j,t} \;+\; \beta_3 * lnIncome_{i,t} \;+\; beta_4 * lnIncome_{j,t} \\ & + \beta_5 * Common\ Language_{i,t} \;+\; \beta_6 * Distance_{i,t} \;+\; \beta_7 * Common\ Border_{i,t} \\ & + \alpha_t + \lambda_{i,j} + \mu_{i,j,t} \end{aligned}$$

where α_t represents year fixed effects, $\lambda_{i,t}$ are dyad fixed effects, and $\mu_{i,j,t}$ is the error term. Note that this model does not include a PTA membership variable. Instead, the gravity model predicts how much states ought to be importing from one another, based on market size and geographic distance as well as other similarities. The residuals from this model are calculated as the difference between the observed trade flow $Imports_{i,j,t}$ and the model prediction. We rely on imports, rather than total trade, because PTAs are defined as agreements that commit members to reciprocal market concessions. Even if the depth of concessions is asymmetric, all countries promise some level of tariff liberalization to one another. Importantly, these promises to liberalize are difficult to keep; they are precisely what makes leaders less likely to honor their predecessors' commitments.

[55] Some economists have criticized the gravity model for not reflecting as much trade as it should because of its inability to capture certain unobserved variables (Trefler, 1995; Anderson and van Wincoop, 2003). However, it remains the standard-bearer for estimating expected levels of trade.

Thus, imports would be the most politically sensitive portion of a country's total trade.

Residuals from this model tell us whether a country is trading at, above, or below expectations for trade, given their economic fundamentals. Residuals of zero would mean that a group of countries traded exactly at the level of what their economic and geographic circumstances would predict, even in the absence of an international agreement. Since it would be to countries' benefit to trade if their fundamentals lined up properly, negative residuals are regarded as representing political barriers to trade.[56] Thus, negative residuals observed in countries that had actually signed a PTA would indicate that they had not removed those political barriers. This makes residuals a very lenient test for implementation; countries that have put an agreement in place need only to trade at or above what their fundamentals would suggest.

Due to the highly skewed nature of bilateral import flows, our outcome $Imports_{i,j,t}$ is logged in the gravity model setup, and therefore our residuals are measured in terms of logged trade values.

The gravity model produces a residual for imports between dyads — that is, each pair of countries i, j. However, we are interested in how changes in domestic leadership affect the implementation of a given PTA p. Therefore, we take the mean of each country i's residuals with all partners in a given PTA p, such that:

$$Implemenation_{i,p,t} \;=\; \overline{Residuals}_{i,j,t}$$

where $Implementation_{i,p,t}$ is country i's average logged residuals across the membership of a given PTA.

There are alternatives to our measure of implementation. The most prominent is looking at the policy reforms each leader enacts. Yet, those specific reforms only matter to the extent they shape trade flows. The distributional outcome we care most about is actual trade, not the underlying policy environment *per se*. As a result, our measure of trade performance provides a useful operationalization of PTA implementation.

Primary Independent Variables. Our core theory relates changes in the domestic political

[56] Rose 2002.

climate to agreement implementation. We argue that when new leaders take office, they are likely to neglect the commitments of the previous leader *when that change is accompanied by a shift in the leader's ideological leanings.*

Measuring ideology — or party platforms, more broadly — is not an easy task.[57] One of the principal challenges is applying any kind of left-to-right spectrum across countries where the meanings of "left" and "right" vary.[58] However, for our purposes, the variation we are interested in is within-country rather than across countries, such that the cross-national validity of the ideology coding is less important than internal consistency. We only need to know whether a given leader has a different political position than the previous leader.[59] Since we can assume that the leader codings are consistent within country, these shifts should be accurately captured by most codings.

The most comprehensive data on ideology is reported in the Database of Political Institutions (DPI). DPI offers a three-value coding, where a value of 1 represents right-wing ideology and a value of 3 represents left-wing. Our core variable of interest is not a leader's ideology, however, but whether this has varied from one leader to the next. We code our indicator $Ideology\ Shift_{i,t}$ as a dichotomous indicator of whether the ideology of the chief executive changed in year t. Note that this measure varies over time. For example, when Labour leader Tony Blair replaces Conservative John Major in 1997, $Ideology\ Shift_{i,t}$ is coded as 1 but it is coded as 0 for the duration of Blair's tenure.

We use additional variables to further isolate the independent effect of ideological shifts. Theoretically, we might expect that moves further to the right, on average, result in policies friendlier to free trade. This stems from a broad association between right-wing ideologies and a commitment to *laissez-faire* economics. We created a more restrictive measure $Right - Wing\ Shift_{i,t}$, which captures whether the ideological shift that took place coincided with a more right-wing leader taking office.

Control Variables and Rival Explanations As discussed in a previous section, a variety of factors may confound our inferences if we failed to control for them. First, there are numerous

[57] Dinas and Gemenis 2010; Lo, Slapin and Proksch 2014.
[58] Benoit and Laver 2006.
[59] This is distinct from whether leader i, t is more or less left-wing, for example, than leader j, t.

features of the domestic political landscape that could potentially effect agreement implementation. To begin with, we want to establish that our mechanism is distinct from turnover generally. It is possible that we would observe a blip in agreement implementation anytime the executive office changes hands *if* the underlying mechanism is just the learning curve associated with new leadership. We therefore include a measure of whether there was $Turnover_{i,t}$ in a given country-year. This variable also captures the general economic conditions that might lead to voter dissatisfaction in a country. If the economy suffered in a leader's tenure, theories of economic voting would predict that an incumbent is more likely to be overturned.

Moreover, one might reasonably argue that the amount of time that leaders have spent in office affects their likelihood of honoring a country's international commitments. Previous work has shown that the duration of a leader's tenure in office has effects in the international system.[60] We include the logged number of days that a leader has been in office to account for whether his or her $Tenure_{i,t}$ also shapes implementation.

We also control for other factors that might separately influence a country's ability to implement the agreements of which it is a member. It is necessary to include these parameters in our estimations to preclude omitted variable bias. For example, leaders might sign PTAs with every intention of implementing them, but poor domestic infrastructure may not be at the levels necessary to fulfill the terms of an international economic agreement.[61] We include a measure of $GDP\ Per\ Capita_{i,t}$ to proxy not only for the overall level of development in a country but also for the size of the market.

Additionally, others might argue that implementation may be driven not by the particular leader in office but rather by the number of political constraints an executive faces; if the number of veto players are high, there may be more opportunities for blocking any kind of implementation.[62] In other words, the ability of new leaders to change policy from the status quo may be conditioned by the structure of their surrounding political institutions. We include the number of $Veto\ Players_{i,t}$ in order to control for the relative ease with which leaders can shape policy. Our information on veto players derives from Henisz's Political Constraints data. It should be noted, however, that

[60] Potter 2007; Zeigler, Pierskalla and Mazumder 2014.
[61] Gray 2014.
[62] Tsebelis 2002.

a high number of veto players would more likely mean that a PTA would not be signed in the first place.[63] That is, political opposition would most likely be most obvious in the bargaining and signing stage and not at the implementation stage.

Implementation may also be derailed if a country suffers from an economic shock. If a country is experiencing a financial crisis, particularly one that involves depreciation of its currency, imported goods will be more expensive and trade may dry up even in the presence of a PTA. We include a measure for $Economic\ Shocks_{i,t}$ by creating a binary variable with a value equal to 1 if in a given year a country's change in consumer prices exceeded a standard deviation of that country's average inflation rate, and 0 otherwise. We also include a measure of *Democracy* to control for the frequency of turnover more generally and also to capture whether democracies are more likely to uphold their commitments, either because of audience costs[64] or because of domestic political institutions.[65] We take this measure from Cheibub, Gandhi and Vreeland 2010.

Aside from domestic factors, the diplomatic relationship between country i and all its partners in PTA p will affect implementation. We include a measure of whether country i has a formal $Alliance_{i,p,t}$ with any of the members of the PTA. Entrenched security ties will reduce fears of exploitation in the international system and ought to encourage economic cooperation.

The strength of the market ties among members also shapes incentives to implement a PTA. We include a measure of $Export\ Share_{i,p,t}$, which measures the proportion of country i's total global exports conducted with the members of PTA p. Higher proportions of exports ought to mean that the PTA is especially important to the domestic economy, increasing incentives to honor the commitment. We also include a measure of market power, measured as $GDP\ Share_{i,p,t}$ — the amount of total PTA GDP accounted for by country i's economy. This control serves a dual purpose. It measures whether country i has a favorable bargaining position relative to its fellow members and it proxies for the overall capacity of the other members. If values of $GDP\ Share_{i,p,t}$ are high, then the legal and bureaucratic capacity of country i is likely to significantly exceed the rest of the membership.

Finally, we include two measures related to the agreement itself. Depth, as distinct from

[63] Mansfield, Milner and Pevehouse 2007.
[64] Tomz 2007.
[65] Leeds 1999; Manger and Pickup 2014.

rigidity, is a measure of the actual trade cooperation $Depth_{i,p}$ that a country agrees to enact in a given PTA. PTAs that ask states to cut tariffs more drastically as well as to include cooperation in trade-related measures such as investment, standards, procurement, intellectual property and competition ought to generate the highest levels of political opposition from the broadest swath of groups at home. We utilize data provided by Baccini, et al. to control for the depth of an agreement. We also include the length of time since the PTA was enacted ($PTA\ Age_{[}p, t]$). Levels of implementation are affected fundamentally by the length of time that governments have had to adjust to their commitments, notwithstanding their willingness to do so.

Our models also include year fixed effects as well as country-PTA fixed effects to account for the unobservable features of time and of specific agreements. In our robustness checks, we introduce a number of other variables that we describe at greater length below.

4 Analysis and Results

We now test the validity of our hypothesis. The results in Table 1 show that ideological turnover is associated with a significant decrease in implementation. This finding is robust to a variety of model specifications and estimation techniques. The results also show that deeper agreements are those that suffer the most from a change in leadership. While deep agreements are sometimes to thought to better tie future leaders' hands, we show that implementation declines significantly under new leaders. Together, the results cast doubt on the idea that agreements are credible commitments.

Baseline Estimates

Hypothesis 1 posits that a change in the ideological leanings of the chief executive results in reduced implementation. Model 1 in Table 1 provides support for our predictions.[66] To begin with, the controls behave largely as expected. Alliance ties and higher numbers of veto players are both

[66] All of the models reported include fixed effects (FEs) for each year t and for each country-PTA i, p. Testing the collective significance of the year dummies shows that they are important to include ($p < 0.001$). Country-PTA fixed effects help us control for unobserved features of the agreement itself, rather than simply including dummies for each country. A Hausman test confirms that FEs are more appropriate than random effects ($p < 0.001$). The standard errors are similarly clustered by country-PTA. These modeling decisions do not affect our results. The baseline models are robust to using just country fixed effects. They are also robust to clustering on countries.

associated with higher levels of implementation. The age of a PTA is also positively associated with implementation, which is consistent with the expectation that, all else equal, countries need time to overcome the costs of adjustment to their trade liberalizing commitments.

Our principal explanatory variable, $Ideology\ Shift_{i,t}$, is significant and negatively signed, suggesting more political barriers to trade when there is a change in leadership. Substantively, the model tells us that a change in ideology results in a 6 percentage point decrease in the residuals. In other words, agreements perform less well when there is a shift in the ideology of the chief executive, by six percentage points (the 95 percent confidence interval around this estimate ranges from -.10 to -.01).[67] This is significant even with the inclusion of a dummy variable for turnover of any type. A change in ideology, not a change in leadership *per se*, has an independent effect on implementation. In substantive terms, holding all other variables at their mean values, a change in the ideology from the incumbent party to a new leader results in a drop in implementation of the European Community's agreement with Iceland (residuals at .04 on average) to its agreement with Morocco (residuals at -.01 on average). It is also equivalent to the average levels of implementation between Hungary's Viktor Orban, a right-wing politician who enacted a program of widespread economic liberalization in that country, and the subsequent leader Peter Medgessy of the socialist party. Thus, the magnitude of the shift has real-world parallels that are of substantive significance.

Note that the direction of the shift is also associated with implementation. Our indicator of whether the ideological shift signaled a move to the right is highly significant and positive, suggesting that transitions toward more right-wing ideologies — typically associated with support for freer trade — result in fuller implementation. This finding is consistent with expectations derived from existing literature in comparative political economy and it gives us increased confidence in our results. However, it is important to note that the inclusion of this measure does not wash out the effect of $Ideology\ Shift_{i,t}$ entirely, suggesting that changes in ideology generally, not just moves toward conservatism, shape implementation.

[67] One should compare the substantive magnitude of these effects with those of other variables that are statistically significant in the models. Holding all other variables at their means, moving GDP share from its minimum to maximum value decreases implementation by -.27 (95 percent confidence interval between -.38 and -.17). The same procedure for export share decreases implementation by -.056 (95 percent confidence interval between -.12 and .001).

4.1 Robustness Checks

We now explore the robustness of our baseline result. To begin with, both for robustness and for ease of interpretation, we directly estimate a gravity equation from a full sample of trade between dyads, from 1970 to 2014. This full sample includes observations between countries that were not members of a particular PTA; thus, the unit of analysis here moves from a monadic one — country/agreement/year — to dyadic. We include in our estimations a variable if the dyads were in a PTA, along with variables capturing whether one of the countries experienced leadership turnover, ideology shift, or a shift to a more conservative government. We then include interactions between the presence of a PTA and all three operationalizations of political change. Using the full sample of countries allows us to observe directly the effects of PTAs on trade; of leadership change on trade; and then the added effect of both a PTA and leadership turnover. This allows the effect of the PTA to vary by leadership change. Table 2 shows these results.

The significance level on the interaction coefficients are difficult to interpret, because they represent whether the difference between both components — the presence of a PTA as well as leadership turnover — is equal to 0. Thus, although these models show that only the switch to a more conservative leader is statistically significant, the F-tests listed below the model results then test whether the interaction is different from each component separately. A significance test implies that there is a difference between the two coefficients. Because the F tests show significance, this indicates that an ideology shift in the presence of a PTA has a significantly smaller effect on trade than a PTA without an ideology shift. There is no difference if there is an ideology shift with a PTA or without. This confirms the baseline results in our initial table.

Durability of Baseline Models

Our next robustness check takes into account the possibility that leadership turnover is non-random: if there exists a generally poor economic climate, that might lead to both low levels of agreement implementation and higher probabilities of leadership turnover. Table 3 addresses this concern by using a two-stage model, in which the first stage estimates the probability of leadership turnover and then the second stage includes that probability into our baseline analysis. Selection models depend on having variables in the first stage that are likely to predict the dependent variable

in the first stage (domestic political change) but are uncorrelated with the outcome in the second stage (PTA implementation). We model the likelihood of turnover using the presence of term limits in a country, the term remaining in a leader's tenure, and economic growth in a country. Although those variables might be linked to turnover as well as to PTA implementation, they are necessary to fully specify the turnover process. Additionally, we also include variables that measure whether the year in question is an election year (taken from Hyde and Marinov 2014), and whether the previous leader died a natural death while in office. These variables ought to predict turnover and an ideological shift while being unrelated to PTA implementation. The Kleibergen-Paap rk Wald F statistic for the test of weak identification falls outside the standard critical values, meaning that concerns of weak instrument bias are alleviated.

The models in Table 3 show these results. The ρ parameter models the correlation of the residuals from the first-stage estimation with those of the second stage; if ρ is significant, it indicates that there remains unmodeled but co-trending processes in both the outcome variable and the variable of interest. Because ρ is insignificant, this concern can be alleviated. Additionally, the two-stage estimation leaves our original result intact: the effect on implementation of an ideology shift remains negative and statistically significant. While this does not fully solve the identification problem, it indicates that once we model the processes that drive turnover, PTA implementation still drops in the presence of an electoral shift.

Next, we split our sample by regime type. While Model 1 controls for whether a state is a democracy, it is likely that our theory operates most effectively when leaders are beholden to a wider constituency. The first two columns of Table 4 report the estimates when splitting the sample — Model 8 for democracies, Model 9 for autocracies. They show that our baseline results are indeed confined to democratic regimes, which also constitute the vast majority of our sample.[68]

We also restricted on EU membership (Table 4). Leaders in each EU member state will vary in their willingness to implement, yet the EU enters into PTAs as a coherent unit. This raises questions about the discretion that individual leaders have in abiding by the larger organization's commitments. We reran Model 1 excluding the EU and find results consistent with the baseline

[68]Splitting the sample roughly approximates interacting $Ideology\ Shift_{i,t}$ with regime type. However, we also introduced an interaction to our baseline model and found it to be highly insignificant statistically and substantively.

estimates (Model 10). We then reran Model 1 only on a sample of EU states and fail to find a significant association between ideological turnover and implementation (Model 11). This is likely because of precisely the issue we raise: individual member states in the EU have certain policies imposed on them, and therefore the hands-tying mechanism appears to work better under a regional organization.

Beyond these sampling restrictions, we included a number of additional variables for potentially confounding factors not soaked up by the fixed effects. One of these is the form of the executive. Whether a government is a presidential or parliamentary system likely conditions the effect of ideological turnover. Executives have more sway in presidential systems than they do in parliamentary systems[69] and would more likely be able to enact policy that would hinder or help trade. In Model 6 (Table 5) we interact $Ideological\ Shift_{i,t}$ with a dichotomous indicator of whether a country has a presidential system. The results show that presidential systems are more likely to implement their commitments to free trade than are their parliamentarian counterparts.

Next, we included measures of each country's recent history (if any) of democratization. Many have argued that democracies both make international commitments more readily and honor those commitments more reliably than do autocracies. Table 6 reports models with controls for the one-, five-, and 10-year changes in each country's Polity score. The results are broadly consistent across these estimations, with only the five-year change narrowly missing conventional levels of statistical significance. Generally, it appears as though a country's transition toward democracy, when it occurs, does not confound our core result.

Table 7 investigates the possibility raised by the rational design literature: the idea that implementation of agreements would be conditioned by agreement design. These arguments would likely predict that deep agreements will be more effective at constraining political behavior, irrespective of ideological turnover. We raised the possibility, by contrast, that deep agreements might actually be *more* likely to be passed over by new leaders, since those agreements might be especially burdensome. Table 7 bears out this argument. On its own, the variable for agreement design does not have statistically significant effects on implementation. However, an interaction between an electoral ideological shift and deep commitments has *negative* effects on implementation that

[69] Lijphart 1999.

are statistically significant. That is, the deeper the agreement, the more likely that an ideological change is likely to lead to lower implementation.

In additional tests omitted for space, we included additional measures of diplomatic ties, including whether there was an ongoing militarized dispute between any of the PTA members and a measure of voting affinity in the United Nations General Assembly. We also looked at different features of domestic political institutions in each country, including whether there are term limits and the length of time that the chief executives party has been in office. The inclusion of these controls did not effect the substantive interpretation of the baseline model.

5 Conclusion

International agreements are intended to be between countries, not leaders. The fundamental assumption behind the logic of agreements as credible commitments is that formal agreements transcend the electoral fates of the politicians who signed them originally. Indeed, the idea that agreements have legacies — that is to say, an agreement's influence extends beyond the tenure of the current leader — is precisely why states may want to formalize the rules of interstate relations. A formal contract should insulate states from changes in their partners' domestic political climates. However, we have demonstrated that these assumptions bear revisiting. An election that results in an ideological shift is associated with low levels of implementation of international economic agreements.

The conditions under which states form PTAs, and how those PTAs are designed, are now reasonably well understood. However, our understanding of membership's effects over time is less complete. In our view, this stems from a failure to appreciate the significant variation in implementation across agreements. PTAs must first be implemented in order for us to make any claims about their perceptible effects on policy. This paper helps specify some of the conditions under which implementation is more or less likely. Our results show that agreements tend to underperform in the wake of leadership turnover and ideological shift. Our results show that treaties do not have monolithic effects; rather, the preferences of leaders — and of their underlying domestic constituents — condition the effectiveness of PTAs.

This offers a more complete picture to the literature on international agreements as credible commitments. Our focus on the consequences of changes in domestic leadership on implementation helps explain why, although very few agreements are formally discarded, leaders do not always enact those agreements. Previous investigations of the effects of IOs on cooperation focus mainly on factors in the interstate system — highlighting the role played by international reputation, for example — or they focus on features of the contract itself. Our argument, however, emphasizes the importance of domestic politics in understanding the cooperative effects over time of international agreements. It shows that the force of international law cannot be divorced from an understanding of the domestic political shifts that could shape the implementation of agreements. While we do not deny the importance of institutional design or interstate politics, we focus chiefly on member governments — the actor ultimately responsible for whether an agreement works in practice — as agents of agreement effectiveness.

Our finding has broad implications for the study of international cooperation. Many studies have lauded the proliferation of international agreements as a sign that states are increasingly willing to defer to international rules. This may be true for the leader who initially signs the agreement. However, we show that subsequent leaders are able to let those agreements slide, even if they do not exit from them formally. This finding casts doubt on the claims about the constraining nature of international agreements over time. Or, at a minimum, it highlights the fundamental importance of domestic politics. The performance of international agreements must be understood in light of domestic political processes. We have offered one cut at that story here, showing that turnover in leaders' preferences leads to less implementation.

The evidence also highlights a particular form of time inconsistency problem — one that raises important questions about institutional design. Previous work emphasizes how economic shocks and other forms of uncertainty about the future state of the world can endanger lasting economic cooperation. Related work argues that PTAs can be designed to cope with this uncertainty. For example, safeguards clauses allow members to erect temporary barriers to market entry. However, designing agreements around changes in the preferences of the member governments is much more vexing. Flexibility provisions, which are typically thought to strengthen agreements by allowing

escape during hard times, are an insufficient solution if the problem is that member states do not want to implement the agreement in the first place. This raises concerns about the ability of agreement design to ensure cooperation even in the event of domestic shocks; leaders may be more likely to simply set an agreement aside than to make explicit use of its designed provision.

These findings also give rise to other questions about the signing of PTAs that have yet to be empirically explored. If new leaders are less likely to implement agreements, we might expect potential PTA partners to be reluctant to forge agreements in the run-up to elections. Other work, however, has argued that PTAs improve leaders' electoral prospects,[70] which would then lead to an expectation that leaders under threat might actually favor signing more PTAs. This proposition would need to be thoroughly investigated in future research.

Future research should examine more closely how states can create more durable agreements. Many have focused on the design of the agreements as critical in ensuring cooperation over time. Our findings, however, demonstrate that deep agreements do not necessarily bind new generations of leaders to those commitments. It is worth investigating the conditions under which the terms of international agreements endure over time, irrespective of who is in office. Moreover, a closer look at governments' preferences, and how those preferences change with domestic political turnover, provides a more thorough understanding of the interaction between domestic politics and international commitments.

[70] Mansfield and Milner 2012; Hollyer and Rosendorff 2012.

References

Abbott, Kenneth and Duncan Snidal. 2000. "Hard and Soft Law in International Governance." *International Organization* 54(3).

Allee, Todd and Jamie Scalera. 2014. "A Beneficial Club, but only if you Pay your Dues: The Effects of GATT/WTO Accession on Trade." Forthcoming, *International Organization*.

Anderson, James E and Eric van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review* 93(1):170–192.

Austen-Smith, David and Jeffrey Banks. 1988. "Elections, Coalitions, and Legislative Outcomes." *American Political Science Review* 82(02):405–422.

Axline, Andrew. 1994. *The Political Economy of Regional Integration: Comparative Case Studies*. Madison, WI: Farleigh Dickinson University Press.

Baccini, Leonardo and Johannes Urpelainen. 2014. "International institutions and domestic politics: can preferential trading agreements help leaders promote economic reform?" *The Journal of Politics* 76(01).

Baier, Scott L., Jeffrey H. Bergstrand and Michael Feng. 2013. "Economic Integration Agreements and the Margins of International Trade." Working paper.

Barnett, Michael and Etel Solingen. 2007. Designed to Fail or Failure of Design? The Origins and Legacy of the Arab League. In *Crafting Regional Cooperation*, ed. Amitav Acharya and Alastair Johnston. Oxford: Oxford University Press pp. 180–220.

Barnett, Michael and Martha Finnemore. 1999. "The Politics, Power, and Pathologies of International Organizations." *International Organization* 53(4):699–732.

Benoit, Kenneth and Michael Laver. 2006. *Party Policy in Modern Democracies*. London: Routledge.

Bergman, Torbjorn, Wolfgang C. Muller, Kaare Strom and Magnus Blomgren. 2003. Democratic Delegation and Accountability: Cross-National Patterns. In *Delegation and Accountability in*

Parliamentary Democracies, ed. Torbjorn Bergman, Wolfgang C. Muller and Kaare Strom. Oxford: Oxford University Press.

Bobick, Talya and Alastair Smith. 2013. "The impact of leader turnover on the onset and the resolution of WTO disputes." *The Review of International Organizations* 8(4):423–445.

Boerzel, Tanja. 2001. "Non-compliance in the European Union: pathology or statistical artefact." *Journal of European Public Policy* 8(5):803–824.

Boerzel, Tanja A. 2005. "Mind the gap! European integration between level and scope." *Journal of European Public Policy* 12(2):217–236.

Bourename, Naceur. 2002. Regional Integration in Africa: Situation and Prospects. In *Regional Integration in Africa*. Paris:OECD pp. 17–43.

Brown, Joseph M. and Johannes Urpelainen. 2014. "Picking Treaties, Picking Winners: International Treaty Negotiations and the Strategic Mobilization of Domestic Interests." Forthcoming, Journal of Conflict Resolution.

Budge, Ian and Richard I. Hofferbert. 1990. "Mandates and Policy Outputs: U.S. Party Platforms and Federal Expenditures." *The American Political Science Review* 84(1):111–131.

Buthe, Tim and Helen V. Milner. 2008. "The Politics of Foreign Direct Investment into Developing Countries: Increasing FDI through International Trade Agreements?" *American Political Science Review* 52:741–762.

Chapman, Terrence and Stephen Chaudoin. 2013. "Ratification Patterns and the Limits of the International Criminal Court." *International Studies Quarterly* 57(2):400–409.

Cheibub, Jose Antonio, Jennifer Gandhi and James Raymond Vreeland. 2010. "Democracy and Dictatorship Revisited." *Public Choice* 143(2-1):67–101.

Colgan, Jeff. 2014. "The Emperor Has No Clothes: The Limits of OPEC in the Global Oil Market.". Forthcoming, *International Organization*.

Dhillon, Amrita and Tomas Sjostrom. 2009. "Leader reputation and default in sovereign debt." Working Paper.

Dinas, Elias and Kostas Gemenis. 2010. "Measuring Parties' Ideological Positions With Manifesto Data: A Critical Evaluation of the Competing Methods." *Party Politics* 16(4):427–450.

Downs, George, David M. Rocke and Peter Barsoom. 1996. "Is the Good News about Compliance Good News for Cooperation?" *International Organization* 50.

Dreher, Axel and Nathan M. Jensen. 2013. "Country or leader? Political change and UN General Assembly voting." *European Journal of Political Economy* 29:183196.

Dür, Andreas, Leonardo Baccini and Manfred Elsig. 2014. "The Design of International Trade Agreements: Introducing a New Database." *Review of International Organizations* .

Eicher, Theo and Christian Henn. 2008. "In Search of WTO Trade Effects: Preferential Trade Agreements Promote Trade Strongly, But Unevenly." University of Washington Working Paper 10/12/08. Seattle, WA; Washington, DC. University of Washington; International Monetary Fund.

Foster, Neil, Johannes Poeschl and Robert Stehrer. 2011. "The impact of Preferential Trade Agreements on the margins of international trade." *Economic Systems* 35(1):8497.

Gray, Julia. 2014. "Domestic Capacity and the Implementation Gap in Regional Trade Agreements." Forthcoming, Comparative Political Studies.

Grossman, Gene M. and Alhanan Helpman. 1995. "The Politics of Free-Trade Agreements." *American Economic Review* 85(4):667–690.

Haftel, Yoram. 2007. "Designing for Peace:Regional Integration Arrangements, Insitutional Variation, and Militarized Interstate Conflcit." *International Organization* 61(1):217–237.

Haftel, Yoram. 2012. *Regional Economic Institutions and Conflict Mitigation: Design, Implementation, and the Promise of Peace.* Ann Arbor, MI: University of Michigan Press.

Haftel, Yoram Z. and Alexander Thompson. 2013. "Delayed Ratification: The Domestic Fate of Bilateral Investment Treaties." *International Organization* 67(2):253–275.

Hicks, Raymond, Helen V. Milner and Dustin Tingley. 2014. "Trade Policy, Economic Interests, and Party Politics in a Developing Country: The Political Economy of CAFTA-DR." *International Studies Quarterly* 58(1):106117.

Hicks, Raymond and Soo Yeon Kim. 2012. "Reciprocal Trade Agreements in Asia: Credible Commitment to Trade Liberalization or Paper Tigers?" *Journal of East Asian Studies* 12(1):1–29.

Hofmann, Tobias and Sooyeon Kim. 2014. "Designing Credibility: The Political Economy of Dispute Settlement Design in Preferential Trade Agreements." Working paper.

Hollyer, James R. and B. Peter Rosendorff. 2012. "Leadership Survival, Regime Type, Policy Uncertainty and PTA Accession." *International Studies Quarterly* 56(4):748–764.

Horowitz, Michael C. and Allan C. Stam. 2014. "How Prior Military Experience Influences The Future Militarized Behavior of Leaders." International Organization. Forthcoming.

Hyde, Susan and Nikolay Marinov. 2014. "Information and Self-Enforcing Democracy: The Role of International Election Observation." *International Organization* 68(2):329–359.

Jensen, Christian B. 2007. "Implementing Europe: A Question of Oversight." *European Union Politics* 8(4):451–477.

Johns, Leslie. 2012. "Courts as Coordinators: Endogenous Enforcement and Jurisdiction in International Adjudication." *Journal of Conflict Resolution* 56(2):257–289.

Kelley, Judith and Jon Pevehouse. 2014. "The Politics of Treaty Ratification." Manuscript - Duke University.

Keohane, Robert. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy.* Princeton, N.J.: Princeton University Press.

Koremenos, Barbara, Charles Lipson and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55(4):761–700.

Leeds, Brett Ashley. 1999. "Domestic Political Institutions, Credible Commitments, and International Cooperation." *American Journal of Political Science* 43(4):979–1002.

Levitz, Philip and Grigore Pop-Eleches. 2010. "Why No Backsliding? The EU's Impact on Democracy and Governance Before and After Accession." *Comparative Political Studies* 43(4):457–485.

Licht, Amanda A. 2010. "Coming into Money: The Impact of Foreign Aid on Leader Survival." *Journal of Conflict Resolution* 54(1):58–87.

Lijphart, Arend. 1999. *Patterns of democracy.* New Haven: Yale University Press.

Lo, James, Jon Slapin and Sven-Oliver Proksch. 2014. "Ideological Clarity in Multi-Party Competition: A New Measure and Test Using Election Manifestos." Working paper.

Mahrenbach, Laura Carsten. 2013. *The Trade Policy of Emerging Powers: Strategic Choices of Brazil and India.* New York, NY: Palgrave McMillan.

Manger, Mark S. and Mark A. Pickup. 2014. "The Coevolution of Trade Agreement Networks and Democracy." Forthcoming, Journal of Conflict Resolution.

Mansfield, Ed, Helen V. Milner and Jon C. Pevehouse. 2007. "Vetoing Cooperation: The Impact of Veto Players on International Trade Agreements." *British Journal of Political Science* 37(03):403–432.

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

Mansfield, Edward and Helen Milner. 2012. *Votes, Vetoes, and the Political Economy of International Trade Agreements.* Princeton, NJ: Princeton University Press.

Mansfield, Edward and Jon Pevehouse. 2006. "Democratization and International Organization." *International Organization* 60(1):137–167.

Marinov, Nikolay. 2005. "Do Economic Sanctions Destabilize Country Leaders?" *American Journal of Political Science* 49(3):564–576.

Martin, Lisa. 2000. *Democratic Commitments: Legislatures and International Cooperation.* Princeton, NJ: Princeton University Press.

McGillivray, Fiona and Alan Stam. 2005. "Political Institutions, Coercive Diplomacy and the Duration of Sanctions." *Journal of Conflict Resolution* 49(5):639–660.

McGillivray, Fiona and Alastair Smith. 2003. "Who can be trusted?: Sovereign debt and the impact of leadership change." Working Paper.

McGillivray, Fiona and Alastair Smith. 2004. "The Impact of Leadership Turnover on Trading Relations Between States." *International Organization* 58(Summer):567–600.

Milner, Helen V and Dustin H Tingley. 2011. "Who Supports Global Economic Engagement? The Sources of Preferences in American Foreign Economic Policy." *International Organization* 65(01):37–68.

Mitchell, Sara McLaughlin and Paul R. Hensel. 2007. "International Institutions and Compliance with Agreements." *American Journal of Political Science* 51(4):721–737.

Moravcsik, Andrew. 2000. "The Origins of International Human Rights Regimes: Democratic Delegation in Postwar Europe." *International Organization* .

Nollen, Stanley D and Dennis P Quinn. 1994. "Free trade, Fair Trade, Strategic Trade, and Protectionism in the US Congress, 1987-88." *International Organization* 48:491–491.

North, Douglass. 1990. *Institutions, Institutional Change and Economic Performance.* New York: Cambridge University Press.

North, Douglass C. and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutional Governing Public Choice in Seventeenth-Century England." *Journal of Economic History* 49(4):803–832.

Peinhardt, Clint and Todd Allee. 2012. "Failure to Deliver: The Investment Effects of US Preferential Economic Agreements." *The World Economy* 35(6):757–783.

Potter, Phil. 2007. "Does Experience Matter?: American Presidential Experience, Age, and International Conflict." *Journal of Conflict Resolution* 51(3):351–378.

Ravenhill, John. 2008. "Fighting irrelevance: an economic community 'with ASEAN characteristics'." *The Pacific Review* 21(4):469–488.

Rose, Andrew K. 2002. "Estimating Protectionism through Residuals from the Gravity Model." Working paper, UC-Berkeley.

Rosendorff, B. Peter and Helen Milner. 2005. "The Optimal Design of International Trade Institutions: Uncertainty and Escape." *International Organization* 55:829–857.

Saunders, Elizabeth. 2011. *Leaders at War: How Presidents Shape Military Interventions*. Ithaca, NY: Cornell University Press.

Schmidt, Manfred. 1996. "When parties matter: A review of the possibilities and limits of partisan influence on public policy." *European Journal of Political Research* 30(2):155183.

Simmons, Beth. 2000. "International Law and State Behavior: Commitment and Compliance in International Monetary Affairs." *American Political Science Review* 94(4):819–835.

Simmons, Beth A. 2001. "The International Politics of Harmonization: The Case of Capital Market Regulation." *International Organization* 55:589–620.

Smith, Alastair. 2009. "Political Groups, Leader Change, and the Pattern of International Cooperation." *Journal of Conflict Resolution* 53(6):853–877.

Smith, Alastair. 2014. "Leader Turnover, Institutions, and Voting at the UN General Assembly." Forthcoming, JCR.

Tallberg, Jonas. 2002. "Paths to Compliance: Enforcement, Management, and the European Union." *International Organization* 56(3):609–643.

Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61(4):821–840.

Torsten Persson, Guido Enrico Tabellini. 2005. *The Economic Effects of Constitutions*. Cambridge, MA: MIT Press.

Trefler, Daniel. 1995. "The Case of the Missing Trade and Other Mysteries." *American Economic Review* 85(5):102946.

Tsebelis, George. 2002. *Veto Players: How Political Institutions Work*. Princeton, NJ: Princeton University Press.

Underdal, Arild. 1998. "Explaining Compliance and Defection: Three Models." *European Journal of International Relations* 4(1):5–30.

Urbatsch, Robert. 2013. "A Referendum on Trade Theory: Voting on Free Trade in Costa Rica." *International Organization* 67(01):197–214.

Vabulas, Felicity. 2014. "Exercising Power or Not? The Role of State Withdrawal from Intergovernmental Organizations." Working power.

Von Stein, Jana. 2005. "Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance." *American Political Science Review* 99:611–622.

Vreeland, James Raymond. 2003. *The IMF and Economic Development*. New York: Cambridge University Press.

Walter, Andrew. 2008. "Governing Finance: East Asia's Adoption of International Standards.".

Wolford, Scott. 2007. "The Turnover Trap: New Leaders, Reputation, and International Conflict." *American Journal of Political Science* 51(4):772788.

Wolford, Scott. 2012. "Incumbents, successors, and crisis bargaining: Leadership turnover as a commitment problem." *Journal of Peace Research* 49(4):517–530.

Zeigler, Sean, Jan H. Pierskalla and Sandeep Mazumder. 2014. "War and the Reelection Motive: Examining the Effect of Term Limits." *Journal of Conflict Resolution* 58(4):658–684.

Table 1: Baseline Estimation

VARIABLES	(1) $Impl_{i,p,t}$
$Ideology\ Shift_{i,t}$	-0.050*
	(0.021)
$Right-wing\ Shift_{i,t}$	0.089**
	(0.027)
$Turnover_{i,t}$	-0.007
	(0.021)
$Tenure_{i,t}$	0.018*
	(0.009)
$Depth_p$	-0.050
	(0.164)
$GDP\ Share_{i,p,t}$	22.802**
	(7.480)
$GDP\ PerCapita_{i,t}$	0.004
	(0.174)
$Democracy_{i,t}$	0.152
	(0.102)
$Veto\ Players_{i,t}$	-0.048
	(0.117)
$Alliance_{i,p,t}$	0.199**
	(0.061)
$Export\ Share_{i,p,t}$	-0.577**
	(0.211)
$Economic\ Shock_{i,t}$	-0.039
	(0.029)
$PTA\ Age_{p,t}$	0.027**
	(0.010)
Constant	-11.341**
	(3.440)
Observations	16,175
Number of state_pta	1,525

Clustered standard errors in parentheses, ** p<0.01, * p<0.05

Table 2: Dyadic Gravity Model

VARIABLES	(2) $LnImp_{i,j,t}$	(3) $LnImp_{i,j,t}$	(4) $LnImp_{i,j,t}$	(5) $LnImp_{i,j,t}$
$GDP_{i,t}$	1.36***	1.36***	0.93***	1.36***
	(0.11)	(0.11)	(0.06)	(0.11)
$Avg.\ GDP_{i,t}$	0.08	0.07	0.19***	0.08
$of\ PTA\ partners_{i,t}$	(0.09)	(0.09)	(0.06)	(0.09)
$Per\ capita\ income_{i,t}$	-0.25**	-0.25**	-0.03	-0.25**
	(0.11)	(0.11)	(0.06)	(0.11)
$Avg.\ Per\ capita\ income$	1.30***	1.31***	1.04***	1.31***
$of\ PTA\ partners_{i,t}$	(0.09)	(0.08)	(0.06)	(0.09)
$Proportion\ of\ GATT\ dyads$	0.10***	0.10***	0.11***	0.10***
$in\ PTA_{i,t}$	(0.02)	(0.02)	(0.02)	(0.02)
$PTA_{i,t}$	0.41***	0.41***	0.40***	0.40***
	(0.03)	(0.03)	(0.03)	(0.03)
$Turnover_{i,t}$	-0.08			
	(0.07)			
$More$	0.08**			
$Conservative_{i,t}$	(0.04)			
$Ideology$	-0.20	-0.07		
$Shift_{i,t}$	(0.14)	(0.11)		
$PTAX$	0.06	-0.04		
$Ideology\ Shift_{i,t}$	(0.14)	(0.11)		
$Turnover_{i,t}$			-0.08	
			(0.05)	
$PTAX$			0.06	
$Turnover_{i,t}$			(0.11)	
$More$				-0.43*
$Conservative_{i,t}$				(0.11)
$PTAX$				0.40*
$Conservative_{i,t}$				(0.24)
$Constant$	-28.68***	-28.45***	-20.21***	-28.66***
	(2.64)	(2.63)	(1.69)	(2.64)
N	282,991	283,603	458,374	282,991
F_{diff}	5.46	14.25	8.83	0.00
p_{diff}	0.02	0.00	0.00	0.98

Table 3: Selection Model

	(6)		(7)	
	Outcome	Selection	Outcome	Selection
VARIABLES	$Impl_{i,p,t}$	$Turnover_{i,t}$	$Impl_{i,p,t}$	$Turnover_{i,t}$
$Ideology\ Shift_{i,t}$	-0.075**		-0.076*	
	(0.027)		(0.027)	
$Right-wing\ Shift_{i,t}$	0.118**		0.119**	
	(0.031)		(0.031)	
$Tenure_{i,t}$	0.018	-0.658**	0.007	-0.656**
	(0.023)	(0.016)	(0.021)	(0.015)
$Depth_p$	0.129**		0.013**	
	(0.018)		(0.018)	
$GDP\ Share_{i,p,t}$	-1.221		-1.214	
	(0.708)		(0.708)	
$GDP\ PerCapita_{i,t}$	-0.077**		-0.076**	
	(0.015)		(0.015)	
$Democracy_{i,t}$	0.240**		0.242**	
	(0.058)		(0.058)	
$Veto\ Players_{i,t}$	0.015	1.961**	0.046	1.941**
	(0.125)	(0.107)	(0.121)	(0.107)
$Alliance_{i,p,t}$	0.084*		0.085*	
	(0.036)		(0.035)	
$Export\ Share_{i,p,t}$	-0.111		-0.111	
	(0.071)		(0.071)	
$Economic\ Shock_{i,t}$	-0.071		-0.070	
	(0.050)		(0.050)	
$PTA\ Age_{p,t}$	0.010**		0.010**	
	(0.002)		(0.002)	
$Term\ Limits_{i,t}$		0.001*		0.001*
		(0.000)		(0.000)
$Years\ Remaining_{i,t}$		0.001**		0.001**
		(0.000)		(0.000)
$GDP\ Growth_{i,t}$		-0.035**		-0.037**
		(0.003)		(0.003)
$Leader\ Natural\ Death_{i,t}$				0.53***
				(.156)
$Election\ Year_{i,t}$				0.30***
				(.003))
Constant	0.948**	4.334**	0.985*	4.269**
	(0.320)	(0.141)	(0.316)	(.141)
Wald Test	0.08		0.74	
Observations	4,123	19,281	4,123	19,281

Clustered standard errors in parentheses, ** p<0.01, * p<0.05

Table 4: Sampling Restrictions

VARIABLES	(8) $Impl_{i,p,t}$ Democracy	(9) $Impl_{i,p,t}$ Autocracy	(10) $Impl_{i,p,t}$ Non-EU	(11) $Impl_{i,p,t}$ EU
$Ideology\ Shift_{i,t}$	-0.047*	-0.049	-0.074*	-0.005
	(0.020)	(0.096)	(0.032)	(0.026)
$Right-wing\ Shift_{i,t}$	0.074**	0.209	0.114**	0.035
	(0.027)	(0.181)	(0.043)	(0.034)
$Turnover_{i,t}$	-0.001	-0.100	-0.052	-0.020
	(0.020)	(0.083)	(0.034)	(0.028)
$Tenure_{i,t}$	0.017*	0.003	-0.007	0.009
	(0.008)	(0.049)	(0.018)	(0.010)
$Depth_p$	-0.007	-0.192	0.133	-0.164
	(0.186)	(0.179)	(0.185)	(0.210)
$GDP\ Share_{i,p,t}$	29.408**	18.824	9.537	32.359*
	(8.592)	(15.362)	(9.074)	(14.721)
$GDP\ PerCapita_{i,t}$	-0.017	0.584	-0.092	0.103
	(0.215)	(0.306)	(0.177)	(0.390)
$Democracy_{i,t}$	0.002	-0.042	0.189	
	(0.093)	(0.166)	(0.103)	
$Veto\ Players_{i,t}$	0.139	-0.230	-0.442*	0.279
	(0.105)	(0.373)	(0.181)	(0.144)
$Alliance_{i,p,t}$	0.234**	0.414	0.071	0.330**
	(0.058)	(0.276)	(0.112)	(0.063)
$Export\ Share_{i,p,t}$	-0.438**	-0.926*	-0.803**	1.013
	(0.154)	(0.468)	(0.229)	(0.585)
$Economic\ Shock_{i,t}$	0.019	-0.166**	-0.069*	0.136
	(0.031)	(0.054)	(0.029)	(0.075)
$PTA\ Age_{p,t}$	0.034**	-0.024	0.040	0.023
	(0.011)	(0.071)	(0.021)	(0.016)
Constant	-14.482**	-13.100	-2.987	-18.047**
	(3.893)	(7.119)	(4.449)	(6.135)
Observations	14,682	1,493	6,190	9,985
Number of state_pta	1,439	169	641	915

Clustered standard errors in parentheses, ** p<0.01, * p<0.05

Table 5: Form of Government: Presidential vs. Parliamentary Systems

VARIABLES	(12) $Impl_{i,p,t}$
$Ideology\ Shift_{i,t}$	-0.041
	(0.021)
$Presidential_{i,t}$	0.626**
	(0.197)
$Pres_{i,t}\ *\ Ideo\ Shift_{i,t}$	-0.071*
	(0.031)
$Right-wing\ Shift_{i,t}$	0.082**
	(0.027)
$Turnover_{i,t}$	-0.009
	(0.020)
$Tenure_{i,t}$	0.015
	(0.008)
$Depth_p$	-0.046
	(0.164)
$GDP\ Share_{i,p,t}$	21.978**
	(7.644)
$GDP\ PerCapita_{i,t}$	0.059
	(0.174)
$Democracy_{i,t}$	0.047
	(0.098)
$Veto\ Players_{i,t}$	-0.020
	(0.116)
$Alliance_{i,p,t}$	0.199**
	(0.062)
$Export\ Share_{i,p,t}$	-0.565**
	(0.211)
$Economic\ Shock_{i,t}$	-0.037
	(0.028)
$PTA\ Age_{p,t}$	0.026**
	(0.010)
Constant	-11.379**
	(3.499)
Observations	16,175
Number of state_pta	1,525

Clustered standard errors in parentheses, ** p<0.01, * p<0.05

Table 6: Regime Change

VARIABLES	(13) $Impl_{i,p,t}$	(14) $Impl_{i,p,t}$	(15) $Impl_{i,p,t}$
$Ideology\ Shift_{i,t}$	-0.049*	-0.035	-0.049*
	(0.021)	(0.020)	(0.021)
$Right-wing\ Shift_{i,t}$	0.091**	0.074**	0.087**
	(0.027)	(0.027)	(0.029)
$Turnover_{i,t}$	-0.009	-0.014	-0.005
	(0.021)	(0.021)	(0.022)
$Tenure_{i,t}$	0.016	0.013	0.017
	(0.009)	(0.009)	(0.009)
$Depth_p$	-0.055	-0.045	-0.109
	(0.165)	(0.166)	(0.175)
$GDP\ Share_{i,p,t}$	23.314**	23.677**	28.570**
	(7.743)	(7.903)	(8.309)
$GDP\ PerCapita_{i,t}$	0.017	0.037	-0.099
	(0.178)	(0.176)	(0.190)
$Veto\ Players_{i,t}$	-0.012	0.018	-0.018
	(0.124)	(0.124)	(0.124)
$Alliance_{i,p,t}$	0.190**	0.195**	0.197**
	(0.061)	(0.059)	(0.066)
$Export\ Share_{i,p,t}$	-0.587**	-0.592**	-0.685**
	(0.211)	(0.211)	(0.216)
$Economic\ Shock_{i,t}$	-0.040	-0.041	-0.055
	(0.030)	(0.030)	(0.031)
$PTA\ Age_{p,t}$	0.028**	0.028**	0.037**
	(0.010)	(0.010)	(0.011)
$Democracy\ Change\ 1yr_{i,t}$	-0.009		
	(0.005)		
$Democracy\ Change\ 5yr_{i,t}$		-0.011*	
		(0.005)	
$Democracy\ Change\ 10yr_{i,t}$			-0.012**
			(0.004)
Constant	-11.671**	-12.406**	-13.767**
	(3.584)	(3.755)	(3.990)
Observations	16,038	15,781	14,437
Number of state_pta	1,521	1,501	1,395

Clustered standard errors in parentheses, ** p<0.01, * p<0.05

Table 7: Agreement Design

VARIABLES	(16) $Impl_{i,p,t}$
$Ideology\ Shift_{i,t}$	-0.044*
	(0.021)
$Depth_p$	-0.049
	(0.165)
$Depth_p\ *\ Ideo\ Shift_{i,p,t}$	-0.050**
	(0.012)
$Right-wing\ Shift_{i,t}$	0.081**
	(0.027)
$Turnover_{i,t}$	-0.004
	(0.020)
$Tenure_{i,t}$	0.018*
	(0.009)
$GDP\ Share_{i,p,t}$	23.757**
	(7.619)
$GDP\ PerCapita_{i,t}$	0.011
	(0.176)
$Veto\ Players_{i,t}$	-0.013
	(0.122)
$Alliance_{i,p,t}$	0.189**
	(0.061)
$Export\ Share_{i,p,t}$	-0.575**
	(0.212)
$Economic\ Shock_{i,t}$	-0.043
	(0.030)
$PTA\ Age_{p,t}$	0.028**
	(0.010)
Constant	-11.812**
	(3.503)
Observations	16,175
Number of state_pta	1,525
R-squared	0.082

Clustered standard errors in parentheses, ** p<0.01, * p<0.05