Stability and change in international policy-making

Magnus Lundgren
Stockholm University

Theresa Squatrito
University of Oslo

Jonas Tallberg
Stockholm University

Thomas Sommerer
Stockholm University

Abstract

International organizations (IOs) have developed into important policy venues beyond the state. Yet our understanding of broader policy dynamics within IOs is limited. This article offers the first systematic study of IO policy agendas. We draw on punctuated equilibrium theory to develop hypotheses about change and stability in IO policy agendas. We examine the policy output of five general-purpose IOs between 1980 and 2015. Statistical analysis of new data comprising nearly 20,000 individual decisions reveals that all five IOs' agendas display punctuated equilibria. We also find that IOs with higher institutional friction, determined by decision rules, membership size, and heterogeneity of membership preferences, display more frequent punctuations. These results suggest that: (1) punctuated equilibrium theory applies to a broader domain than previously thought; (2) patterns of change in IOs are more complex than conventionally expected; and (3) institutional friction matters for IOs' responsiveness to changing societal demands and global problems.

Key words: policy agenda, punctuated equilibrium, international organizations, institutional performance

Word count: 9,606

Over recent decades, international organizations (IOs) have developed into important policy venues beyond the nation state. Not only have IOs grown in numbers, but more importantly, these organizations have been vested with more far-reaching policy-making authority in an increasing range of domains (Hooghe and Marks 2015). While some of these IOs are task-specific organizations that specialize in particular policy domains, such as the International Atomic Energy Association (IAEA), others are general-purpose organizations that cover the full range of problems confronting a set of member states, such as the European Union (EU) (Lenz et al. 2015). Part of this expansion of international authority has come at the expense of national policy-making, while another part responds to new problems not previously dealt with domestically. Accordingly, it has become common to speak of "global health policy," "global trade policy," and "global environmental policy."

Yet, despite this development, we know little about the broader dynamics of IO policy-making. Importantly, how responsive is IO policy-making to changing societal demands and pressing global problems? Are policy agendas marked by stability and inertia or by openness to change, as new issues and ideas produce shifts in political attention? When change occurs, is this a smooth and seamless process of adjustment or a disruptive and dramatic break with the status quo? What determines whether IOs are capable of flexibly adapting their policy agendas to new information? These are issues of great importance for our understanding of policy-making in IOs, but which so far have received limited systematic attention. To gain traction on these matters, this article engages in the first comparative analysis of macro patterns in IO policy-making. We ask two specific questions. First, what dynamics of stability and change characterize the policy agendas of IOs? Second, what explains variation in these dynamics across IOs and over time?

Theoretically, we turn to the study of policy agendas in American and Comparative Public Policy for inspiration. The most powerful expectation in this area originates from punctuated equilibrium theory (PET), originally developed by Baumgartner and Jones (1993) and subsequently refined by others (Baumgartner, Green-Pedersen, and Jones 2006; Jones and Baumgartner 2005; Jones, Sulkin, and Larsen 2003). According to PET, political attention is scarce and institutions are marred by friction, leading to policy agendas characterized by periods of stability, when there is little or no change, and periodic punctuations, marked by dramatic and rapid change. Empirical research in American and Comparative Public Policy has consistently shown that policy dynamics tend to follow the distinct pattern predicted by PET. In International

Relations, the notion of punctuated equilibrium has informed analyses of institutional change (Colgan et al. 2011; Krasner 1984; Young 2010). However, PET's expectations about stability and change in policy agendas have never been systematically tested on a broad range of IOs.[1]

Empirically, this article examines the policy output of five general-purpose IOs between 1980 and 2015, based on newly gathered data on nearly 20,000 individual policy decisions. The IOs are the African Union (AU), EU, Organization of American States (OAS), Organization of Islamic Cooperation (OIC), and United Nations (UN). The IOs have been selected to ensure wide geographical scope, a focus on general-purpose organizations, and comparable policy output. The empirical analysis maps the overall patterns in IO policy attention, evaluates the presence of policy punctuations, and assesses explanations of variation in punctuation across the five IOs.

The analysis reveals two principal findings. First, all five IOs have policy agendas that reflect the pattern of punctuated equilibria predicted by PET. Frequency distributions of year-on-year policy changes exhibit the presence of agenda punctuations: most of the time, agendas remain stable, but occasionally, there are dramatic shifts in the attention awarded a particular policy area. Second, this pattern is more pronounced when and where IO decision-making is associated with higher institutional friction, in terms of demanding decision rules and large memberships with heterogeneous preferences. When institutional friction is higher, flexibly adapting policy agendas to information about changing societal demands and problem pressures is more difficult, resulting in periods of stasis interrupted by occasional dramatic change. Conversely, when institutional friction is lower, smoothly adjusting policy agendas to new information is easier, leading to smooth and continuous change and fewer disruptive large-scale shifts.

These results suggest three broader implications for research on public policy and global governance, further developed in the conclusion. First, they demonstrate that PET's scope of applicability is broader than previously thought, and the differences between national and international policy-making fewer than conventionally believed. While repeated studies have shown that PET predicts policy dynamics in diverse national contexts, we are the first to show its broad applicability beyond its traditional home domain. Related, these findings challenge the notion of international policy-making being qualitatively different from national policy-making

[1] To our knowledge, the only existing study is Alexandrova et al. 2012, which tests (and finds support for) PET in the context of the EU.

because of an absence of hierarchical authority in the international realm. Second, the patterns we reveal suggest a more nuanced understanding of IOs' propensity for change. While it has been common in International Relations to emphasize why IOs should be resistant to change, we join a growing group of scholars identifying and explaining varying patterns of change in global governance (for example Rixen et al. 2016). The central question should not be whether IOs are amendable to policy change – they clearly are – but how that change comes about and what it looks like. Third, the findings speak to an emerging literature on the performance of IOs (Gutner and Thompson 2010; Tallberg et al. 2016), suggesting that IOs are more likely to be responsive to changing societal demands and global problems when institutional friction is low. This finding is significant since the institutional design of IOs, in terms of decision rules and membership size, is a choice within the control of member states (Koremenos et al. 2001).

The Argument: Punctuated Equilibrium Theory and International Policy-Making

We suggest that PET offers a fruitful approach for understanding the policy dynamics of IOs. In the following, we first outline the general tenets of PET, then explain why it is potentially applicable to IOs, and finally identify specific hypotheses for empirical testing.

Baumgartner and Jones (1993) developed PET to explain the tendency of policy agendas to be stable most of the time with interruptions of periodic abrupt change. PET builds on the assumption that governmental attention is scarce. Governing entities have a limited capacity and cannot attend to all policy problems. Moreover, the problems that do receive attention do so to varying levels. Over time, agendas will change in a punctuated pattern as policy-makers respond irregularly to new information and attention shifts from some issues to others (Jones and Baumgartner 2005). Long periods of inattention will be supplanted by short periods of attentiveness due to shifts in factors such as framing, social mobilization, or venue control (Baumgartner, Green-Pedersen, and Jones 2006, 961). This punctuation pattern contrasts with the image of smoothly functioning organizations, in which policy agendas change gradually as new information about the state of the world is seamlessly integrated.

Empirically, this dynamic of punctuated equilibria has been established in several studies on American public policy, comparative public policy (for example, Baumgartner et al. 2009;

Breunig 2006), and more recently EU public policy (Alexandrova, Carammia, and Timmermans 2012). In addition, punctuated equilibria have been observed in a range of different policy fields, including public health policy (Baumgartner and Jones 1993), environmental policy (Repetto 2006), and welfare policy (Jensen 2009).

Despite the extensive literature on policy agendas, policy punctuations have not been systematically studied in the international domain. Yet there are good reasons to expect that PET would wield predictive power in this setting as well. The central conditions contributing to punctuated equilibrium in the national context are present in the international domain as well. First, just like the national policy venues for which PET was originally formulated, IOs are policy venues, institutions authorized to make decisions on issues (Baumgartner and Jones 1993). This characterization applies to IOs, and increasingly so, since IOs have been conferred growing authority over recent decades (Hooghe and Marks 2015). As Barnett and Finnemore state, IOs "make authoritative decisions that reach every corner of the globe and affect areas as public as governmental spending and as private as reproductive rights" (2004, 1).

Second, much like national political systems, IOs are prone to capacity limitations (Barnett and Finnemore 1999). IO policy-makers have cognitive limitations resembling bounded rationality (Jupille, Mattli, and Snidal 2013; Poulsen 2014) and IOs typically experience attention scarcity, in the sense that the number of issues deserving political attention exceed the capacity of policy-makers to develop rules and norms. This goes particularly for general-purpose IOs, which share the encompassing policy scope of national political systems (Lenz et al. 2015). This leads us to expect that IOs, too, will devote unequal attention to issues and that issues will rise or fall on agendas in response to, for instance, social mobilization.

Third, anecdotal evidence suggests that global policy-making is marked by patterns of punctuated equilibria. For example, most observers of world affairs can recall moments when public attention to particular issues spiked, from the global financial crisis in 2008 to the more recent refugee crisis in 2015 and the Zika virus in early 2016. These shifts in attention were observed among IOs as well, as the International Monetary Fund, UN, and World Health Organization struggled to respond to these policy challenges. Moving beyond anecdotal evidence, a recent study establishes that the EU's policy agenda follows a pattern of punctuated equilibria (Alexandrova, Carammia, and Timmermans 2012).

Assuming that PET is applicable to global policy-making as well, it generates two core propositions – one descriptive and one explanatory. First, PET posits that one of the key dynamics of policy agendas is that they follow patterns of punctuation where stability alternates with radical change. The cognitive limitations of policy-makers and capacity limitations of institutions cause policy agendas to tend toward relative stability. Stability is occasionally broken when attention shifts from one issue to another due to a large-scale event, a previously ignored problem receiving attention, or pressure on a policy reaching a threshold level (Princen 2013, 855). Setting aside the causes leading to individual punctuations, the overall descriptive pattern posited by PET is that of long periods of stability in policy agendas punctuated by rapid change. We hypothesize:

H1: The policy agendas of IOs display patterns of punctuation, combining longer periods of stability with shorter periods of radical change.

Second, PET offers a potential explanation of variation across IOs and over time in the degree of policy punctuation. PET submits that institutional friction contributes to the degree of punctuation across organizational contexts. Institutional friction can be understood as the costs involved in the translation of policy inputs into policy outputs (Jones, Sulkin, and Larsen 2003). In other words, institutional friction is the hurdles to policy change in a specific context. As Princen explains: "To the extent that policy-making institutions impose such friction, policy change will be less frequent (leading to longer periods of stability) but once it comes the 'correction' will be greater (leading to larger punctuations)" (2013, 858). Research in Comparative Public Policy confirms that institutional friction reinforces the degree of punctuation in a policy agenda (Baumgartner et al. 2009; Jones and Baumgartner 2005).

Institutional friction of IOs is likely to be generated by a combination of three factors: decision rules, membership size, and preference heterogeneity. This expectation is firmly anchored in rational institutionalist scholarship on domestic (Scharpf 1997; Tsebelis 2002) and international (Lake and Powell 1999) policy processes. To begin with, IOs vary in their institutional rules on decision-making (Blake and Payton 2015). Decision rules can contribute to friction, affecting the ability to arrive at policy decisions (Scharpf 1997; Tsebelis and Yataganas 2002). IOs with unanimity rules, in comparison to those with majority voting and weighted

voting, confront greater hurdles to decision-making. For example, both the UN Security Council and the WTO are IOs where unanimity rules raise barriers to flexible decision-making.

In addition, membership size serves as a source of institutional friction. The number of actors is widely understood as shaping the likelihood and nature of international cooperation (Axelrod and Keohane 1985; Koremenos, Lipson, and Snidal 2001). For IOs, the number of member states can likewise affect the ease of decision-making. All else equal, the transaction costs for arriving at a decision in a large multilateral setting with more than 100 member states, like the WTO, are larger than those in a small minilateral setting with few member states, such as the Nordic Council of Ministers. For this reason, larger membership sizes likely increase institutional friction and thus the degree of punctuation in IOs' policy output.

Finally, preference heterogeneity among member states contributes to institutional friction. Previous research has shown that the distribution of preferences among member states affects institutional design, institutional change, and delegation in international politics (Axelrod and Keohane 1985; Hawkins et al. 2006; Koremenos, Lipson, and Snidal 2001). When member states' preferences diverge more, agreement on new or revised policies will be more difficult to achieve. Examples include the UN Security Council, where rapprochement between the permanent members after the end of the Cold War led to a noticeable change in policy output (Malone 2004), and the EU, where the preference heterogeneity among its member states has affected the speed of decision-making (König 2007). All else equal, the more divergent member state preferences are, the more difficult it is to reach decisions, and the more likely we are to see a pattern of policy punctuations.

Accordingly, we arrive at a second hypothesis:

H2: Policy agendas of IOs subject to greater institutional friction will be characterized by a higher degree of punctuation.

Our hypotheses draw attention to descriptive (H1) and explanatory (H2) patterns in IO policy dynamics. If PET applies to IOs, as per H1, we would expect to see patterns of policy punctuations among IOs. Empirically, policy punctuations are observed by studying period-to-period policy change distributions. Our confidence in H1 would be increased if we observe that IO policy change distributions deviate from the normal distribution in a specific pattern known as

leptokurtosis, marked by "tall peaks" and "long tails." If PET can explain variation in punctuation, as per H2, we would expect to see that IOs with higher institutional friction are systematically different from IOs with lower institutional friction. Empirically, high-friction IOs are then expected to exhibit period-to-period change distributions that are more peaked and have longer tails than IOs with lower degrees of institutional friction.

Data on IO Policy Agendas

To test these hypotheses, we gathered separate datasets on the policy agendas of five IOs: AU, EU, OIC, OAS, and UN. This sample of IOs provides a suitable starting point for the comparative analysis of international policy-making for three reasons. First, it has a wide geographical scope, including the paramount global IO, the UN, and key regional IOs on all major continents. Second, it exclusively includes general-purpose IOs with a broad policy agenda, as opposed to task-specific IOs that are focused on narrow set of issues (Lenz et al. 2015). These IOs all deal with a range of issues, from security to development and human rights. General-purpose IOs are the most appropriate for evaluating PET in the international setting, partly because a broad policy scope is required to study shifts in political attention across issue areas, and partly because these IOs are those that most approximate the all-purpose orientation of domestic political systems. Third, the policy output of these IOs has a comparable format and style. All adopt decisions separated by issue and sector, as compared with IOs that adopt decisions in an aggregate manner and present them to the world in the form of communiqués. In short, this is a sample of geographically representative IOs with a multi-issue orientation producing comparable policy output.

For each IO, we gathered data on key policy outputs (acts) from the principal intergovernmental decision-making body with IO-wide membership in the period from 1980 to 2015, as per Table 1.[2] The choice of bodies ensures that the policy output emerges at a

[2] The UN offers a choice between two key intergovernmental decision-making bodies, the General Assembly and the Security Council. We decided on the former since the Security Council is not a general-purpose body, making it unsuitable for agenda analysis, and has a membership that is restricted to a subset of IO members, making it less comparable to the other IOs in the sample. Our calculations show that the results presented below hold regardless of which UN body is selected (see online appendix).

comparable stage of the policy process (see Baumgartner et al. 2009). In total, the dataset contains 18,523 policy acts with a total length exceeding 100,000 pages.

Table 1: *Overview of data sources*

IO	Years	Policy acts	N
African Union[3]	1980-2014	Assembly Decisions, Resolutions	846
European Union	1980-2013	Council Directives	2,148
Organization of American States	1980-2014	General Assembly Resolutions	2,161
Organization of Islamic Cooperation	1980-2015	Council Resolutions	2,851
United Nations	1980-2015	General Assembly Resolutions	10,517

To identify the allocation of attention across each IO's policy agenda, we first assigned a topic code to each policy act based on computer-assisted content analysis. We employed a keyword topic classifier (Grimmer and Stewart 2013; Laver and Garry 2000) and a list of topic codes for the main policy areas in global governance (see Appendix 1). The general structure of our policy content categories mirrors that of the codebook of the Comparative Agendas Project (CAP) (for an overview, see John 2006), dividing the policy space into a set of general categories that are subsequently divided into more specialized policy domains, each with its own sub-code.[4] The general categories cover the key policy concerns in global governance, including economic development, human rights, and security, but also certain domestic concerns that are subject to frequent international regulation and coordination, such as health policy or labor policy.

We assigned topics to individual policy acts based on the relative frequency of key words drawn from newly constructed topic-specific dictionaries for each of our 16 main topic codes. In

[3] Our dataset for the AU includes the predecessor organization, the Organization of African Unity (OAU). The OAU transformed into the AU in 2002. At this time, the principal policy output type changed from Decisions to Resolutions, which is why we include both in the analysis.

[4] Since the domestic orientation of CAP policy content categories make them unsuitable for our purposes, we created a new list, geared towards the key concerns of IOs. This allowed us to expand the range of categories available to global governance policy topics and reduce the range of categories exclusive to the domestic sphere.

constructing the dictionaries, we analyzed existing policy research dictionaries[5], policy-specific literature in global governance, and IO policy documentation. To increase specificity, the dictionary for each topic contains both 1- and 2-grams. For example, the dictionary for the "Security and defense" category contains both "conflict" and "conflict prevention," and the dictionary for the "Environment and natural resource management" lists both "climate" and "climate change".

In accordance with standard practice, we prepared documents for quantitative analysis by carrying out a series of text transformations, including stemming (the removal of word ends), tokenization (turning the text into a "bag of words"), and the removal of punctuation, capitalization, and very common words (so called "stop words").[6] For each of the five IOs, we then classified the processed documents by topic, assigning each document the topic code that best corresponded to the content based on keyword frequency. Aggregating across years, we then calculated the distribution of agenda space devoted to a single policy issue, defined as the proportion of policy acts assigned to that topic in relation to the aggregate policy production in the given year.

We evaluated the reliability of the classification algorithm by comparing it to handcoding for the EU, OAS, OIC, and AU (OAU). For each of these, we randomly selected five years and handcoded all its output for those years (between 247 and 400 documents per IO), assigning topic codes based on manual semantic analysis. When compared with the computer algorithm, handcoding yields overlapping (within 5 percentage points) proportion estimates for between 69 and 88 percent of the categories, depending on IO. The average overlap is 78 percent.

Results

We begin by reporting our findings on the overall patterns in IO agendas, before assessing the presence of policy punctuations (H1) and explanations for variation in punctuation across IOs

[5] Extant keyword dictionaries for agenda research have generally been developed for the analysis of domestic policy (Grimmer and Stewart 2013) or for the EU (Alexandrova, Carammia, and Timmermans 2012), and consequently do not contain a sufficient range of topic codes to support a comparative analysis of global policymaking.

[6] All automatized text analysis was carried out using the tm package in R (version 3.2.3). Scripts and dictionaries are available from the authors.

(H2). We conclude by reporting findings from a longitudinal analysis of the five IOs, including three case illustrations.

IO agenda patterns

Table 2 and Figure 1 show the allocation of policy attention across the agendas of the five IOs. Several IOs share common features in the distribution of attention across topics. First, all IOs have fairly concentrated agendas, where a few topics dominate the policy output. Across the IOs, the top five topics account for between 66 percent (EU) and 83 percent (UN) of the agenda. Second, there is a significant amount of overlap in agenda orientation across IOs. The top concerns of global governance – economic development, human rights and security – are found among the major topics on all IO agendas except that of the EU. Reflecting the deeper integration of its member states, the EU has an agenda oriented slightly more toward domestic policy areas, such as energy and transport, science, and labor. Many IOs spend a considerable amount of their attention on questions of IO governance, which include administration, budgeting, and membership issues. The OIC, however, only devotes two percent of its agenda to IO governance, which may reflect its lower degree of institutionalization. Third, all IOs have agendas that, besides the top items, cover a broad range of minor topics, suggestive of the "long tail" that has been identified for policy agendas in the domestic domain (Breeman et al. 2009). Overall, these are patterns we would expect from general-purpose IOs with a wide policy remit.

Table 2: Distribution of policy topics on IO agendas, 1980-2015

Policy topic	AU	EU	OAS	OIC	UN
IO governance	12%	4%	9%	2%	13%
Economic development	21%	2%	23%	13%	19%
Health and social affairs	5%	5%	2%	1%	1%
Human rights	11%	6%	29%	12%	26%
Culture and education	3%	1%	4%	14%	1%
Labor and employment	11%	8%	1%	2%	0%
Law and crime	1%	1%	7%	1%	3%
Governance	1%	3%	2%	0%	1%
Environment	3%	13%	2%	1%	2%
Agriculture and commodities	2%	7%	1%	1%	1%

Energy and transport	2%	19%	1%	1%	1%
Science and technology	3%	8%	2%	1%	2%
Trade and industry	3%	3%	4%	4%	1%
Finance and monetary	5%	17%	4%	6%	2%
Security	12%	0%	10%	20%	22%
International affairs	1%	0%	1%	20%	3%
Top five topics	*68%*	*66%*	*77%*	*78%*	*83%*
TOTAL N	798	2,091	2,397	2,851	10,517

Within each IO, there is considerable temporal variation both in the volume of policy output and the share of attention devoted to individual topics (Figure 1). The data are suggestive of volume punctuations in all IOs. The OAS and EU exhibit signs of short-term changes followed by a return to long-term average volume levels, and the UN and OIC exhibit marked shifts followed by relative stability around new equilibria. In the case of the OIC, the scope of the overall agenda displays a trajectory suggestive of ebb and flow, with a marked peak in 2003 followed by a period of lower policy production. Similar surges can be seen in the AU, around the time the OAU transitioned into the new organizational structure of the AU, and the UN, from the mid-1980s.

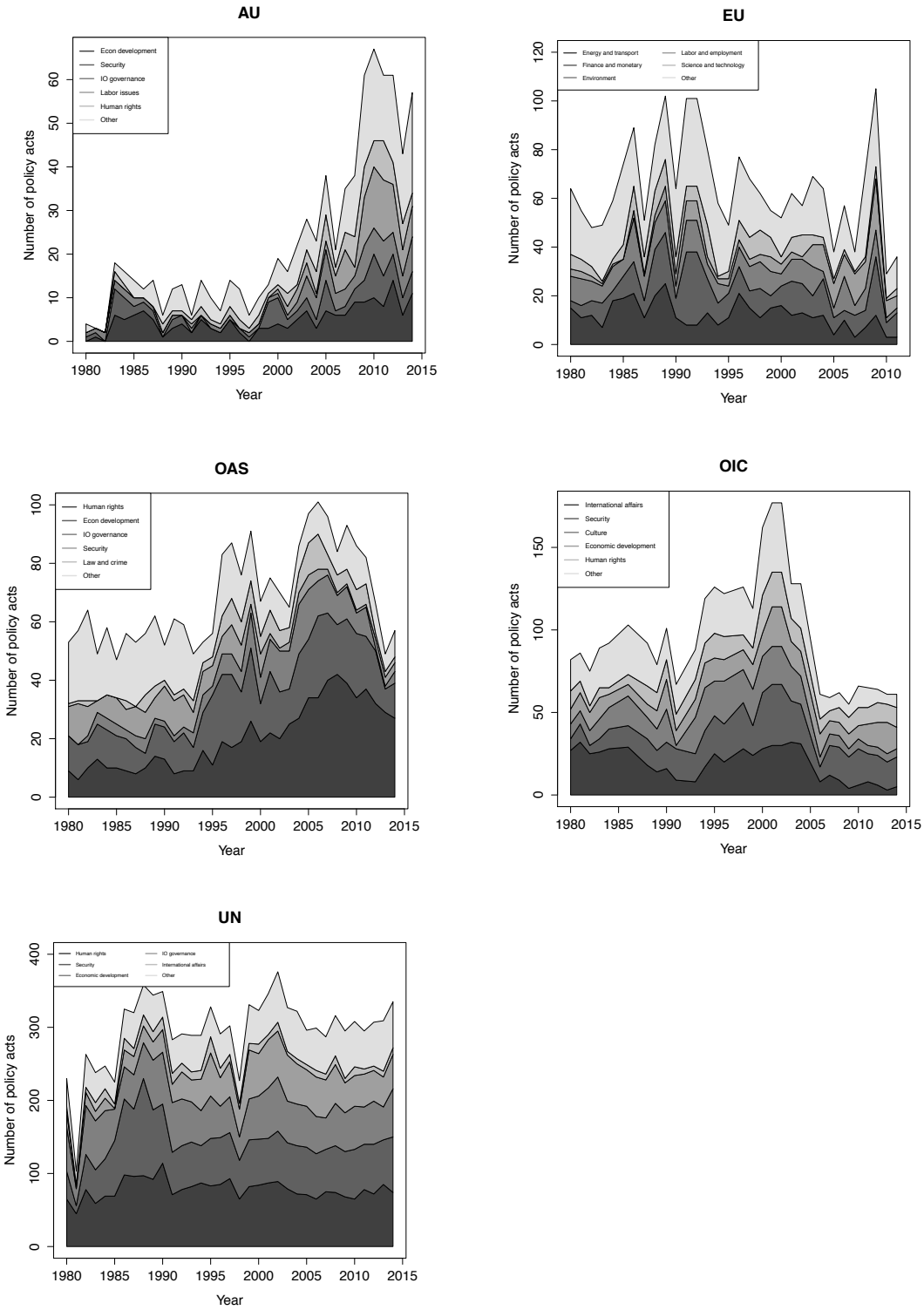

Figure 1: *Agenda allocation, 1980-2015*

With respect to agenda attention, there is evidence of both stability and change in each of the five IOs. All IOs display periods of attention stability, in which the allocation of output across topics remains proportional, even though the overall volume of the agenda expands or contracts. These periods of relative agenda stability include, for example, the OIC in the 1980s and the EU between 1995 and 2005. Change, however, is frequent. The data contain both signs of gradual change, such as the steady rise in attention awarded by the OAS and UN to the topic of human rights, and of radical change, such as the dramatic increase in AU's attention to labor issues after 2005, and the near disappearance of security from the agenda of the OAS.

In sum, it is evident that all IOs have agendas that are subject to a considerable amount of fluctuation over time, both in terms of overall volume and attention to specific topics. These are patterns that point to the possible presence of policy punctuations. To ascertain whether this is the case, we carry out further tests.

Policy punctuations in IO agendas

The conventional method to identify the presence of agenda punctuations is stochastic process analysis of population-level data. Rather than engaging in point prediction of individual punctuations, this approach identifies punctuations based on the characteristics of frequency plots of annual changes in agenda attention (Jones 2005; Jones et al. 2009). To generate year-to-year change distributions, we calculated the percentage-percentage shift in agenda attention allocated to each topic for each year between 1980 and 2015. To avoid artificially induced dramatic changes, we include only the five most frequent topic codes for each IO, grouping outstanding items under an "Other" category.[8]

We evaluate the presence of policy punctuations via frequency distribution plots and distributional shape statistics. Figure 2 presents histograms of year-on-year changes, with the magnitude of change on the x-axis and the count of observations on the y-axis. For purposes of comparison, we superimpose a normal curve with a mean and standard deviation identical to the given distribution. In a policy process without friction, we would expect a distribution of period-to-period changes to overlap with the normal curve, as policy responds flexibly to changes in the state of the world. In contrast, for a policy process characterized by institutional friction and punctuation, we would expect a distribution that is leptokurtic; that is, a distribution with a tall peak and thicker tails.

[8] Our results are robust with respect to different methods of calculating change, including percentage-count methods (see online appendix).

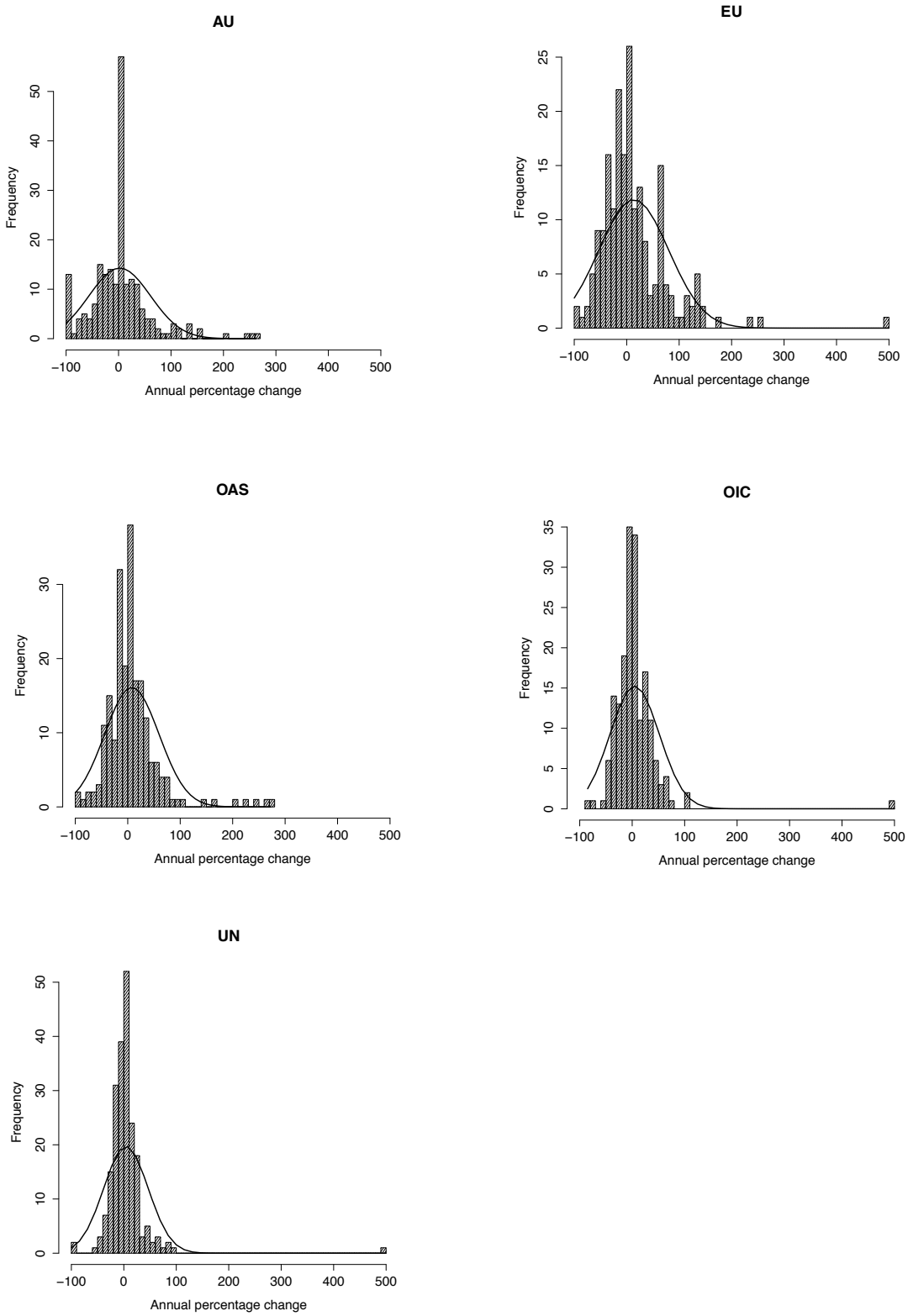

Figure 2: *Year-on-year percentage agenda changes, 1980-2015*

The change distributions in the data on the five IOs suggest general leptokurtic tendencies. For all IOs, we observe a high number of observations around zero, which

corresponds to periods in which the attention awarded to a topic remained stable. These observations reflect an IO policy agenda in equilibrium. However, the distributions also contain a higher than expected (under the normal distribution) frequency of observations at either extreme, including some observations that are very far off on the right tail.[10] These observations signify dramatic changes in the allocation of attention to a particular issue: punctuations of the equilibrium. For example, a score of +200 percent indicates a tripling of attention, whereas a score of -100 percent indicates that a topic has been dropped from the agenda completely. While changes of this magnitude are rarely seen in punctuation studies focused on budgetary allocations (Jones et al. 2009), it frequently happens in studies of policy agendas, which tend to be more volatile (Alexandrova, Carammia, and Timmermans 2012).

The observed patterns of slender peaks and extended, thicker tails suggest that the distributions are leptokurtic. This finding is further borne out by distribution shape statistics, presented in Table 3. We include statistics for both kurtosis and L-kurtosis, as both are commonly used in the literature, but we attach more weight to the L-kurtosis scores, which are robust to smaller sample sizes (Royston 1992).

Table 3: *Shape statistics for frequency distributions of year-on-year changes in policy attention, 1980-2015.*

IO	N	Kurtosis	L-kurtosis	Skew	Mean	SD	S-W test	K-S test
AU	210	13.5	0.305	1.20	0.07	0.63	0.00	0.00
EU	204	14.1	0.228	2.64	0.14	0.67	0.00	0.00
OAS	210	8.47	0.301	3.73	0.08	0.52	0.00	0.00
OIC	216	29.7	0.345	6.53	0.05	0.47	0.00	0.00
UN	216	99.2	0.355	8.43	0.04	0.41	0.00	0.00

A normal distribution has a kurtosis of 3 and a L-kurtosis of 0.123. Higher values mean that a distribution is leptokurtic (high peak and thick tails); lower values that it is platykurtic (low peak and thin tails). All of the five IOs exhibit kurtosis and L-kurtosis scores well above what would be expected had the data been generated from a normal distribution. The smallest L-kurtosis value, 0.228 for the EU, is already a significant deviation from 0.123, and the largest value, 0.355 for the UN, suggests marked leptokurtic properties. Non-normality is further confirmed by Shapiro-Wilk and Kolmogorov-Smirnov tests for the five IOs, all of which comfortably reject the null hypothesis of normal distributions.

[10] We group changes larger than 500 percent in the 500 percent category.

To place our results in the context of previous scholarship, we compare the L-kurtosis statistics with Baumgartner et al. (2009), who report that the change distributions for parliamentary bills, laws, and statutes in Denmark, Belgium, and the United States have L-kurtosis scores of 0.26, 0.29, and 0.25, respectively. These are domestic policy outputs that emerge at a stage in the policy process that is comparable to the adoption of resolutions and directives by IOs. We note that the magnitude of our kurtosis scores is somewhat higher, on average, but overall in line with the previous results for comparable output distributions at the domestic level.

In sum, the evidence presented above is congruent with what we would expect from the logic of policy punctuations. We observe annual change distributions with tall peaks and thick tails, as expected for agendas that change little for long periods of time but occasionally are subject to dramatic change. Hence, the data are strongly indicative of the presence of punctuated equilibria in these IOs (H1).

Comparative analysis

The data suggest that there is variation across the five IOs with respect to the degree of punctuation. The distribution of the EU exhibited lower leptokurtosis than any of the other IOs, whereas the UN had the highest score. What could explain this variation? Previous research has demonstrated a correlation between institutional friction and leptokurtic distributions in outputs (Baumgartner et al 2009; Jones et al 2009) and we proceed to test if the same association can be found in our data. While institutional friction conceptually includes both cognitive limitations and institutional barriers (Jones and Baumgartner 2005), actual tests have focused on the latter (Jones et al. 2009). We follow that practice.

Above, we identified three sources of institutional friction in IOs: decision rules, membership size, and preference heterogeneity. We do not claim to have identified a final and exhaustive list of factors that drive institutional friction in IOs, but we think that these factors provide a reasonable starting point. *Decision rules* is a measure of the degree to which member state preferences must align before a decision can be adopted, as manifested in IO voting rules. The lowest barriers are raised in IOs that employ simple majority voting, whereas the most restrictive rules require unanimity, that is, all member states must agree. *Membership size* is the count of IO member states. All else equal, a higher number of member states means higher transaction costs in decision-making. *Preference heterogeneity* is an ordinal composite measure that seeks to tap into underlying interest heterogeneity stemming

from differences among the core characteristics of IO members, such as developing versus developed economies, large versus small states, and autocracies versus democracies.

Following a procedure used by Jones et al (2009), we measure and rank each IO on these three variables, to produce an additive institutional friction index (Table 4). The ordinal rank of each IO for each variable is presented in parentheses. For decision-making, we took into consideration the proportion of votes required for adopting a decision in relation to the membership as a whole, and the extent to which more flexible decision-making processes (e.g., shifting from a requirement of unanimity to majority voting) could be invoked (Tsebelis and Yataganas 2002). For the size of membership, we relied on data from the Correlates of War IGO dataset (Pevehouse, Warnke, and Nordstrom 2004) and updated by IO-specific research as necessary. We derive a measure of preference heterogeneity via a deductive method, seeing preferences as determined by a basket of observable member state characteristics (see Frieden 1999: 61ff), including income level, economic size, political regime, population, and cultural composition. This approach rests on the assumption that countries that vary along these factors are likely to hold different views on core matters discussed within IOs, such as economic development (e.g., redistribution between rich and poor countries), security (e.g., the relative contribution of small and large countries in the provision of security), and normative concerns (e.g., negotiations on human rights between liberal democracies and authoritarian regimes). The greater the heterogeneity among member states along these five observable dimensions, the greater the deducted preference heterogeneity. [11] Accordingly, preference heterogeneity scores were calculated from normalized variance statistics on population, overall size of the economy (GDP in the year 2000), level of development (GDP per capita in the year 2000), level of democracy (Polity IV 2015), and level of ethno-linguistic fractionalization (Fearon 2003).[12]

Table 4: *Indicators of institutional friction in IOs*

IO	Decision rules	Membership size	Preference heterogeneity	Total friction
AU (OAU)	Consensus (alt. 2/3 majority) (3)	54 (3)	9 (2)	8
EU	Qualified majority (2)	28 (1)	8 (1)	4

[11] In our case, measures of direct preferences such as voting records in the UN General Assembly (Voeten et al. 2009) are not suitable because they are endogenous to the outcome variable.
[12] See Appendix 2 for data.

OAS	Simple majority (alt. 2/3 majority) (1)	35 (2)	12.5 (3)	6
OIC	Consensus (alt. 2/3 majority) (3)	57 (4)	13.5 (4)	11
UN	Formally 2/3 majority or simple majority, but in practice consensus on most decisions. (3)	193 (5)	19.5 (5)	13

The results suggest that there is a significant amount of variation in institutional friction across IOs. Of the five organizations, the EU has the lowest institutional friction, primarily a function of a smaller and comparatively homogeneous membership. The UN, the largest and most heterogeneous organization, is associated with the highest institutional friction.

To evaluate the presence of a correlation between institutional friction and leptokurtosis, we plot the two variables in Figure 4. Our sample size is very limited, but the plot suggests a correlation between institutional friction, as approximated by our ordinal ranking, and L-kurtosis statistics. The best-fitting line has a positive slope.[13] The conclusion is that institutional arrangements that inject friction into decision-making processes are associated with a higher likelihood of punctuations. Hence, based on these limited data, the prediction in hypothesis 2 is borne out.

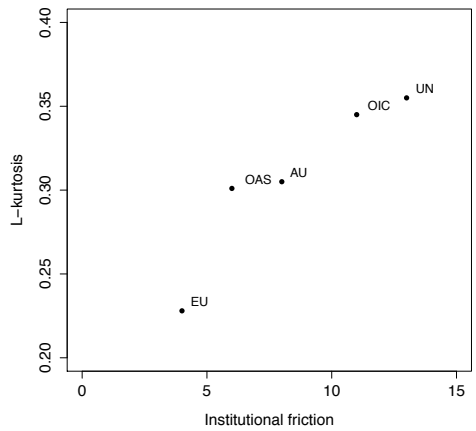

Figure 4: *The relationship between institutional friction and L-kurtosis.*

[13] The correlation coefficient, r, is 0.93. Regressing leptokurtosis scores on institutional friction yields a linear model, $y = 0.20 + 0.013x$, where y is leptokurtosis and x institutional friction. The coefficient for x is statistically significant at the 95% level.

It is possible that our additive measure obscures underlying variation. To evaluate the impact of the different factors included in the additive friction index, we plot the values of the three variables *decision rules*, *membership size*, and *preference heterogeneity* against the leptokurtosis values in Figure 5. This more fine-grained analysis reveals that all three factors are positively correlated with leptokurtosis values. The correlation is least distinct with respect to decision rules, where OAS is an outlier and three organizations receive the same ordinal score. For membership size, the correlation is relatively clear: IOs with more members display a greater tendency to policy punctuations, as measured by leptokurtosis statistics. Likewise, the plot for preference heterogeneity exhibits high degrees of correlation between the variables, suggesting that IOs with more diverse membership preferences are more likely to have agendas that evolve through punctuations.

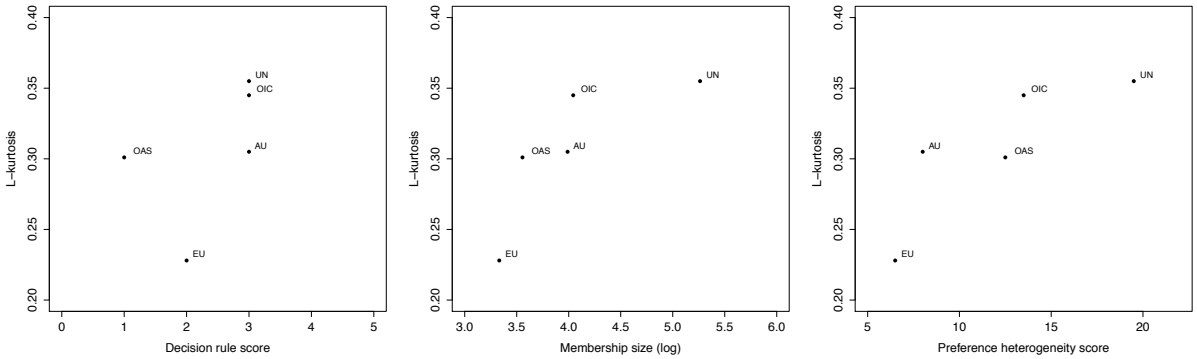

Figure 5: *The relationship between individual measures of institutional friction and L-kurtosis.*

The data is thus congruent with the hypothesized association between institutional friction and policy punctuations. Both visual inspection of cross-plots and regression statistics confirm that IOs with more institutional friction tend to have higher degrees of punctuation, as measured by leptokurtosis in change distributions.

Longitudinal analysis

To further evaluate the robustness of the link between institutional friction and agenda patterns, we measure the degree of punctuation before and after distinctive IO-specific shifts in institutional friction. Such longitudinal comparisons allow us to get at the effect of shifts in the variables of interest, while holding much of the organizational context constant.

Table 5 presents an overview of this analysis, detailing the year and relevant context of the shift in institutional friction and the resulting change in L-kurtosis. These results are interesting in two main respects. First, while the cutoffs vary between the IOs, it is clear that there is no common trend in IO policy-making: some IOs exhibit accentuating punctuation levels, while others have moved toward lower levels, and one, the EU, has remained largely stable. In other words, viewed as a group, IOs have not changed consistently in terms of policy punctuations, suggesting that variation is due to IO-specific factors and not an exogenous factor that leads IOs to over time develop smoother and more efficient policy-making. Second, shifts in agenda patterns are intelligible in light of changes in institutional friction. The general tendency for IOs is to exhibit changes in L-kurtosis that match what would be predicted based on information about institutional friction. Perhaps the clearest example is the UN, which exhibits a severely disjointed agenda during the years of Cold War confrontation, but a considerably lower rate of punctuation in the years after 1991, suggesting that greater preference alignment ushered in an era of more efficient policy-making. Similarly, the 2002 transitioning of the failing OAU into a new organization, the AU, is associated with a significant reduction of L-kurtosis. The institutional shift is thus not only associated with a significant increase in policy output (see Figure 1), but also a smoother translation of inputs into outputs. In the following, we look in detail at the other three IOs in our sample, investigating how changes in decision rules (EU), preference heterogeneity (OAS), and membership size (OIC) affect the degree of punctuation, while also highlighting interactions with other components of institutional friction.

Table 5: *Longitudinal analysis*

IO	Shift point	L-kurtosis before shift	L-kurtosis after shift	Context
AU	2002	0.39	0.20	As the OAU was transformed into the AU in 2002, an improved institutional structure was established.
EU	1999	0.23	0.22	The Treaty of Amsterdam entered into force, increasing the remit of qualified majority voting, in parallel with continued expansion of the EU membership.
OAS	1994	0.28	0.32	The establishment of the Summit of the Americas widened the agenda, opening up new preference fault lines.
OIC	1991	0.21	0.41	The end of the Cold War led to an increase in membership as several Muslim majority Post-Soviet states acceded the OIC.
UN	1991	0.46	0.22	The end of the Cold War caused greater preference homogeneity.

In the case of the EU, the constant level of leptokurtosis over time is consistent with developments in its institutional friction. While the EU has changed greatly in two of the three components of institutional friction during our observation period, these changes have tended to cancel each other out. Beginning with the entry into force of the Single European Act in 1987, the EU has gradually shifted from unanimity to qualified majority voting as the decision-making rule in the Council. With every new treaty since then (Maastricht Treaty 1993, Amsterdam Treaty 1999, Nice Treaty 2003, and Lisbon Treaty 2009), the EU has brought more policy domains into the remit of qualified majority voting, making that the standard principle of present-day decision-making. Research demonstrates that this change in the decision rule has had a clear positive effect on decision-making efficiency in the EU (Golub 1999; Schulz and König 2000; Golub and Steunenberg 2007). However, during the same time period, the EU's membership has more than tripled, from 9 in 1980 to 28 in 2013. This development has affected the institutional friction in the opposite direction. In fact, the reason why the EU has shifted toward qualified majority voting is partly to counteract a potential negative impact on its decision-making efficiency from an ever growing number of member states. While the third component of institutional friction, preference heterogeneity, has remained relatively constant, a development exogenous to our measure of institutional friction – the growing involvement of the European Parliament in EU policy-making – is known to have affected decision-making capacity in a negative direction as well (Golub 1999; Golub and Steunenberg 2007; Schulz and König 2000). We use the Maastricht Treaty as cut-off point in Table 5, but given the gradual nature of these changes, we also repeated the analysis with the other treaties as cut-off points. The results remained the same: a constant level of leptokurtosis over time (see online appendix).

In the case of the OAS, we observe a higher rate of punctuations after 1994 (Table 5). Although two components of institutional friction remained relatively constant during the observation period, with a minor increase in member states and no major change in the decision rules, we find evidence that changes in preference heterogeneity can be linked to the shift in L-kurtosis. More specifically, the OAS membership has become more democratic, decreasing the preference heterogeneity of the IO. Processes of democratization in member states such as Brazil (1985) and Chile (1989) transformed the OAS from an IO with an almost equal share of authoritarian and democratic members in 1980 into a predominantly democratic one by the end of the Cold War. This change is mirrored in the sharp decrease of the variance of Polity scores among OAS members around that time, but also in a growing number of human rights policies on the OAS agenda thereafter (Figure 1). The downward

trend in institutional friction would lead to the expectation of a lower rate of dramatic and disruptive changes on the agenda. Does this mean that the case of the OAS contradicts our theoretical expectation? A closer look at the OAS tells another story. What looks as diminishing heterogeneity due to regime change at first glance turns out to be the opposite. Several authoritarian regimes in Latin America were supported by the US government during the 1980s (for example Forsythe 1995), and the end of these regimes did not lead to greater alignment with the United States. This cleavage grew during the 1990s, when countries like Venezuela, Bolivia, Brazil and others swerved left (Castañeda 2006). Another development beyond the limits of our narrower operationalization of institutional friction also contributed to greater preference heterogeneity among member states. The first 'Summit of the Americas' in Miami in 1994 marked a new orientation of the OAS toward issues of regional economic integration and the Free Trade Area of the Americas (FTAA) (Feinberg 2006; Horwitz 2011, 29). The issue of trade liberalization led to greater tensions within the OAS, as manifested in strong opposition against the FTAA and the failure of negotiations in 2003 (Kennedy 2003; Prevost 2005).

The OIC represents a second case of growing policy punctuation over time. As Table 5 shows, the L-kurtosis score almost doubled after 1991.[15] This development corresponds to a higher degree of institutional friction after the end of the Cold War. The OIC membership grew by almost one third between 1980 and 2000. The accession during the 1990s of seven post-Communist countries, two Latin American countries, and two countries from Sub-Saharan Africa transformed the original OIC from an organization with a largely Middle Eastern membership into a more global organization. While the decision-making rule of qualified majority voting for resolutions in the Council of Ministers remained unchanged, growing preference heterogeneity from the early 1990s onwards also contributed to a higher level of institutional friction (Kayaoglu 2015). While tensions between US-aligned member states and Soviet bloc member states existed during the Cold War, the polarization within the OIC grew in the wake of crises like the Iraqi invasion of Kuwait in 1990 (Akbarzadeh and Connor 2005, 82). What followed was greater rivalry between OIC leader states like Iran, Turkey, Egypt, and Pakistan, and a declining role of Saudi Arabia that previously dominated and financed the organization (Akbarzadeh and Connor 2005; Kayaoglu 2015, 41).

[15] A prominent pattern in the policy volume of the OIC is the steep decrease of the number of resolutions in 2003 (Figure 1). We used this year as an alternative shift point to corroborate our analysis and received similar results of increasing L-kurtosis over time.

Conclusion

This article has offered the first systematic multi-N analysis of stability and change in IO policy-making. Theoretically, we have drawn on PET, as originally developed in the study of policy agendas in American and Comparative Public Policy. Empirically, we have mapped and explained agenda dynamics in five general-purpose IOs between 1980 and 2015.

Two principal findings emerge from our analysis. First, IO policy agendas exhibited patterns consistent with punctuated equilibria. Of the five IOs examined, none displayed a normal output distribution, which is the expectation if agendas adapt flexibly to inputs. Rather, they were all characterized by longer periods of stability, interrupted by shorter periods of dramatic departure from the equilibrium.

Second, while the IO policy agendas were invariably punctuated, some agendas exhibited more punctuation than others. We established that institutional friction was an important determinant of the degree of punctuation in IO policy making. Comparing across the sample, the IOs with the highest level of institutional friction – UN and OIC – were those with the most punctuated policy agendas, while the IOs with the lowest institutional friction – EU and OAS – had agendas characterized by less punctuation. Longitudinal analysis demonstrated how changes in institutional friction within the IOs, holding other organizational characteristics constant, translated into the expected effects on patterns of punctuation.

Our findings have three broader implications for the study of public policy and global governance. First, the predictive power of PET is broader than previously thought. With only one exception (Alexandrova, Carammia, and Timmermans 2012), previous scholarship in this tradition has focused on the domestic level, identifying punctuations in the agendas of national and local government. This article represents the first comparative application of PET to IO policy-making and demonstrates that policy punctuations are equally prominent in global governance. This finding also suggests that claims about qualitative differences between national and international decision-making are exaggerated (Mearsheimer 1994; Nielson and Tierney 2003). Despite international policy-making taking place through collective decisions among composite actors in an absence of hierarchical authority, IOs produce policy dynamics very similar to those of national legislatures and municipal councils.

Second, IOs are more amenable to change than conventionally expected in global governance research. How that change comes about and the forms it takes should be the focus for future research. Existing research tends to emphasize the reasons why IOs are resistant to

change, citing overly demanding decision hurdles (Scharpf 1988), stabilizing organizational cultures (Barnett and Finnemore 2004), and feedback effects that lock in the status quo (Fioretos 2011). Our results suggest a more nuanced picture and a reorientation of the agenda toward studying mechanisms and patterns of change, mirroring recent developments in historical institutionalist research (Rixen et al. 2016; Streeck and Thelen 2005).

Third, the findings have implications for our understanding of IO performance – the subject of a growing literature in recent years (Gutner and Thompson 2010; Hafner-Burton, von Stein, and Gartzke 2008; Tallberg et al. 2016). We show that IOs are more likely to be more responsive to changes in the state of the world, such as societal demands and global problem pressures, when institutional friction is lower. This is a scientifically and politically important finding. It does not mean that there is a quick fix to IO performance, since states sometimes design IOs not to be efficient and since producing policy only is a first step toward impact. But it suggests that the basic parameters of decision-making – the actors, their preferences, and the rules that guide their interaction – are fundamental to IO performance and partly within the power of states to affect.

Bibliography

Akbarzadeh, Shahram, and Kylie Connor. 2005. "The Organization of the Islamic Conference: Sharing an illusion." *Middle East Policy* 12(2): 79-92.

Alexandrova, Petya, Marcello Carammia, and Arco Timmermans. 2012. "Policy Punctuations and Issue Diversity on the European Council Agenda." *Policy Studies Journal* 40(1): 69-88.

Axelrod, Robert, and Robert O. Keohane. 1985. "Achieving Cooperation under Anarchy: Strategies and Institutions." *World Politics* 38(1): 226-254.

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

Barnett, Michael N., and Martha Finnemore. 2004. *Rules for the World : International Organizations in Global Politics*, *Rules for the World*. Ithaca, NY: Cornell University Press.

Baumgartner, Frank R., Christian Breunig, Christoffer Green-Pedersen, Bryan D. Jones, Peter B. Mortensen, Michiel Nuytemans, and Stefaan Walgrave. 2009. "Punctuated Equilibrium in Comparative Perspective." *American Journal of Political Science* 53(3): 603-620.

Baumgartner, Frank R., Christoffer Green-Pedersen, and Bryan D. Jones. 2006. "Comparative Studies of Policy Agendas." *Journal of European Public Policy* 13(7): 959-974.

Baumgartner, Frank R., and Bryan D. Jones. 1993. *Agendas and Instability in American Politics*. Chicago: University of Chicago Press.

Blake, Daniel J, and Autumn Lockwood Payton. 2015. "Balancing Design Objectives: Analyzing New Data on Voting Rules in Intergovernmental Organizations." *The Review of International Organizations* 10(3): 377-402.

Breeman, Gerard, David Lowery, Caelesta Poppelaars, Sandra L. Resodihardjo, Arco Timmermans, and Jouke De Vries. 2009. "Political Attention in a Coalition System: Analysing Queen's Speeches in the Netherlands 1945–2007." *Acta Politica* 44(1): 1-27.

Breunig, Christian. 2006. "The More Things Change, the More Things Stay the Same: A Comparative Analysis of Budget Punctuations." *Journal of European Public Policy* 13(7): 1069-1085.

Castañeda, Jorge G. 2006. "Latin America's Left Turn." Foreign Affairs 85 (May/June): 28.

Colgan, Jeff D., Robert O. Keohane, and Thijs Van de Graaf. 2011. "Punctuated Equilibrium in the Energy Regime Complex." *Review of International Organizations* 7(2): 117-143.

Copelovitch, Mark S. 2010. "Master or Servant? Common Agency and the Political Economy of IMF Lending." *International Studies Quarterly* 54(1): 49-77.

Fearon, James. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8(2): 195-222.

Feinberg, Richard E. 2006. "Presidential Mandates and Ministerial Institutions: Summitry of the Americas, the Organization of American States (OAS) and the Inter-American Development Bank (IDB)." *Review of International Organizations* 1(1): 69-94.

Fioretos, Orfeo. 2011. "Historical Institutionalism in International Relations." *International Organization* 65(2): 367-399.

Forsythe, David P. 1995. "Human Rights and US Foreign Policy: Two Levels, Two Worlds." *Political Studies* 43(1): 111-130.

Golub, Jonathan. 1999. "In the Shadow of the Vote? Decision Making in the European Community." *International Organization* 53(4): 733-764.

Golub, Jonathan, and Bernard Steunenberg. 2007. "Forum Section: How Time Affects EU Decision-Making." *European Union Politics* 8(4): 555-566.

Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21(3): 267-297.

Gutner, Tamar, and Alexander Thompson. 2010. "The Politics of IO Performance: A Framework." *Review of International Organizations* 5(3): 227-248.

Hafner-Burton, Emilie M., Jana von Stein, and Erik Gartzke. 2008. "International Organizations Count." *Journal of Conflict Resolution* 52(2): 175-188.

Hawkins, Darren G., David A. Lake, Daniel L. Nielson, and Michael J. Tierney. 2006. *Delegation and Agency in International Organizations*. Cambridge: Cambridge : Cambridge University Press.

Hooghe, Liesbet, and Gary Marks, 2015. "Delegation and Pooling in International Organizations." *Review of International Organizations* 10 (3):305–328.

Horwitz, Betty. 2011.*The Transformation of the Organization of American States: A Multilateral Framework for Regional Governance*. London: Anthem Press.

König, Thomas. 2007. "Divergence or Convergence? From Ever-growing to Ever-slowing European Legislative Decision Making." *European Journal of Political Research* 46: 417-444.

Jensen, Carsten. 2009. "Policy Punctuations in Mature Welfare States." *Journal of Public Policy* 29(3): 287-303.

John, Peter. 2006. "The Policy Agendas Project: A Review." *Journal of European Public Policy* 13(7): 975-986.

Jones, Bryan D. 2005. "A Model of Choice for Public Policy." *Journal of Public Administration Research and Theory* 15(3): 325-351.
Jones, Bryan D., and Frank R. Baumgartner. 2005. *The Politics of Attention: How Government Prioritizes Problems*. Chicago: Chicago University Press.
Jones, Bryan D., Frank R. Baumgartner, Christian Breunig, Christopher Wlezien, Stuart Soroka, Martial Foucault, Abel François, Christoffer Green-Pedersen, Chris Koski, Peter John, Peter B. Mortensen, Frédéric Varone, and Stefaan Walgrave. 2009. "A General Empirical Law of Public Budgets: A Comparative Analysis." *American Journal of Political Science* 53(4): 855-873.
Jones, Bryan D., Tracy Sulkin, and Heather A. Larsen. 2003. "Policy Punctuations in American Political Institutions." *American Political Science Review* 97(1): 151-169.
Jupille, Joseph, Walter Mattli, and Duncan Snidal. 2013. *Institutional Choice and Global Commerce*. Cambridge: Cambridge University Press.
Kayaoglu, Turan. 2015. *The Organization of Islamic Cooperation: Politics, Problems, and Potential*. London: Routledge.
Kennedy, Kevin C. 2003. "FTAA Negotiations: A Melodrama in Five Acts." International Law Review 1(2): 121-138.
Koremenos, Barbara, Charles Lipson, and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55(4): 761-799.
Lake, David A. and Robert Powell. 1999. *Strategic Choice and International Cooperation*. Princeton, NJ: Princeton University Press.
Laver, Michael, and John Garry. 2000. "Estimating Policy Positions from Political Texts." *American Journal of Political Science* 44(3): 619-634.
Lenz, Tobias, Jeanine Bezuijen, Liesbet Hooghe, and Gary Marks. 2015. "Patterns of International Authority: Task-Specific vs. General-Purpose Organizations." *Politische Vierteljahresschrift* 49(1-2): 131-153.
Malone, David. 2004. *The UN Security Council: from the Cold War to the 21st century*. London: Lynne Rienner Publishers.
Mearsheimer, John J. 1994/1995. "The False Promise of International Institutions." *International Security* 19(3): 5-49.
Moravcsik, Andrew 1997. "Taking Preferences Seriously: A Liberal Theory of International Politics." *International Organization* 51(4): 513-553.
Nielson, Daniel and Michael Tierney. 2003. "Delegation to International Organizations: Agency Theory and World Bank Environmental Reform." *International Organization* 57(1): 241-276.
Pevehouse, Jon, Kevin Warnke, and Timothy Nordstrom. 2004. "The Correlates of War 2 International Governmental Organizations Data Version 2.0." *Conflict Management and Peace Science* 21(2): 101-119.
Polity IV. 2015. "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2010." http://www.systemicpeace.org/inscrdata.html.
Poulsen, Lauge N. 2014. "Bounded Rationality and the Diffusion of Modern Investment Treaties." *International Studies Quarterly* 58(1): 1-14.
Prevost, Gary. 2005. "Contesting Free Trade The Development of the Anti-FTAA Movement in the Streets and in the Corridors of State Power." *Journal of Developing Societies* 21(3-4): 369-387.
Princen, Sebastiaan. 2013. "Punctuated Equilibrium Theory and the European Union." *Journal of European Public Policy* 20(6): 854-870.
Repetto, Robert, ed. 2006. *Punctuated Equilibrium and the Dynamics of U.S. Environmental Policy*. New Haven: Yale University Press.

Rixen, Thomas, Lora Viola, and Michael Zürn, eds. 2016. *Historical Institutionalism and International Relations: Explaining Institutional Development in World Politics*. Oxford: Oxford University Press.
Royston, Patrick. 1992. "Which Measures of Skewness and Kurtosis Are Best?" *Statistics in Medicine* 11(3): 333-343.
Scharpf, Fritz W. 1997. *Games Real Actors Play : Actor-Centered Institutionalism in Policy Research*, *Games Real Actors Play*. Boulder, CO: Westview Press.
Schulz, Heiner, and Thomas König. 2000. "Institutional Reform and Decision-making Efficiency in the European Union." *American Journal of Political Science* 44(4): 653-666.
Tallberg, Jonas, Thomas Sommerer, Theresa Squatrito, and Magnus Lundgren. 2016. "The Performance of International Organizations: A Policy Output Approach." *Journal of European Public Policy* 23(7): 1077-1096.
Tsebelis, George. 2002. *Veto Players: How Political Institutions Work*. Princeton, NJ: Princeton University Press.
Tsebelis, George, and Xenophon Yataganas. 2002. "Veto Players and Decision-Making in the EU after Nice." *Journal of Common Market Studies* 40(2): 283-307.
Voeten, Erik, Strezhnev, Anton, and Bailey, Michael, 2009, "United Nations General Assembly Voting Data", hdl:1902.1/12379, *Harvard Dataverse*, V16 [UNF:6:V09q7uLWRKrY16w2g/cmDg==]
Young, Oran R. 2010. *Institutional Dynamics : Emergent Patterns in International Environmental Governance*, *Earth System Governance*. Cambridge, MA: MIT Press.

Appendix 1: Topic codes

Code	Topic
100	IO governance
200	Economic development
300	Health and Social Affairs
400	Human rights
500	Culture and education
600	Labor and employment
700	Law and crime
800	Governance
900	Environment and natural resource management
1000	Agriculture, fisheries, commodities
1100	Energy and transport
1200	Science and technology
1300	Trade, economic integration, and industry
1400	Finance and monetary policy
1500	Security and defence
1600	International affairs

Appendix 2: Preference heterogeneity data

We calculated preference heterogeneity scores based on variance statistics on five dimensions: level of development (GDP per capita in the year 2000), level of democracy (Polity IV 2015; ranging from -10 to 10, with higher values signifying a higher level of democracy), size of the economy (GDP in the year 2000), population (in millions in the year 2000), and level of ethno-linguistic fractionalization (ELF) (Fearon 2003; ranging from 0 to 1, where higher values signify greater cultural heterogeneity). We scored each IO's value based on its ordinal ranking (1-5) and aggregated the ordinal rankings across the five dimensions to reach a heterogeneity score, which was itself ranked. (Since population and size of economy are strongly correlated, they are weighted at 50% in the aggregation.)

Preference heterogeneity *(Standard deviation of member state values for each dimension)*						
IO	**Economic development** (SD, GDP/cap)	**Democracy** (SD, Polity score)	**Population** (SD, millions)	**Size of economy** (SD, GDP)	**Culture** (SD, ELF)	**Preference heterogeneity score**
AU (OAU)	2,426 (1)	4.96 (3)	21.6 (1)	58.6 (1)	0.26 (3)	**9 (2)**
EU	4,558 (2)	0.28 (1)	22.6 (2)	646 (3)	0.17 (1)	**8 (1)**
OAS	9,414 (4)	2.50 (2)	65.4 (4)	2,290 (5)	0.24 (2)	**12.5 (3)**
OIC	7,966 (3)	5.47 (4)	45.0 (3)	171 (2)	0.27 (4)	**13.5 (4)**
UN	11,440 (5)	6.35 (5)	151 (5)	1,182 (4)	0.29 (5)	**19.5 (5)**