States choosing to sign treaties seem to be heavily influenced by the simultaneous signing decisions of other states. However, using the number of other signers as an independent variable in statistical tests can violate the assumption of independence of observations. We show, using Beth Simmons’ 2000 *APSR* article “International Law and State Behavior: Commitment and Compliance in International Monetary Affairs”, that tests which neglect this “simultaneity dilemma” produce non-robust results regarding regional diffusion. We correct for Simmons’ diffusion variables by using lagged data and find that previous regional choices have no statistically or substantively significant effect on current choices to join Article VIII of the IMF Agreement. We also demonstrate that using linear functional forms for diffusion covariates may be misleading. We provide evidence from a number of treaties that various nonlinear functional forms are more appropriate for estimating differing diffusion processes.
1. Introduction

What affects states’ decisions over whether to join treaties? Other than the usual set of cost/benefit explanations – state interests, treaty incentives, fear of punishment – one particularly intuitive factor is the joining decisions of other states. On the assumption that treaties produce either network effects (they become more valuable the more members they have), or negative externalities (countries that are not party to a treaty face hitherto inexistent costs because others join), we can hypothesize that states will face monotonically increasing incentives to join treaties as other states join. Empirical evidence supporting this hypothesis could have important analytical and policy implications – suggesting both methods for controlling recalcitrant states and for ensuring that international agreements are more effective.

However, producing confirmatory tests of such hypotheses is no simple task. Because states are likely to wait for other states to join before committing to a treaty, they face a coordination dilemma similar to the famous “Stag Hunt” game. While a region of states might be unequivocally better off if they all join a treaty together, it may be the case that being the sole joiner in a region produces a particularly undesirable outcome. In this scenario, states need to coordinate their treaty accession to avoid such a “sucker” outcome; thus treaty-joining decisions are very likely to be simultaneous rather than incremental.

This analytical puzzle is simple to analyze using game-theoretical tools and yet extremely problematic for statistical analysis. Because states’ decisions to join treaty regimes are interdependent, in cross-sectional time-series analysis we cannot control for

---

1 Even in cases where the costs for sole joiners do not appear to be dramatic, a region of states might nonetheless want to coordinate their joining decision because of transactions costs or so as to share implementation costs.
the number of other states who choose to join in a given year if we wish to examine influences on a particular state’s joining decisions. Put simply, this would mean using the measures of the dependent variable for all states in a given year as an independent variable affecting each state’s individual decision in that same year. The analysis would require committing one of the cardinal sins of social analysis: analyzing change in the dependent variable by using results from that very same dependent variable as an independent variable. The only statistically appropriate way of measuring the effect of other states’ joining decisions requires the use of lagged measures of the proportion of previous joiners – thereby absolving us of the endogeneity sin. In this article we demonstrate this “simultaneity dilemma” using Beth Simmons’ “International Law and State Behavior: Commitment and Compliance in International Monetary Affairs”.\(^2\) We find that the ‘regional effect’ described in the article\(^3\) vanishes once lags are used to correct for the simultaneity bias.

Another consideration one should take into account when analyzing states’ treaty-joining decisions is the overall pattern of diffusion. If the joining decisions of other states are significant predictors of individual treaty accession, we should also expect the rate of diffusion to have a noticeable effect. Analytically this indicates that the functional form of diffusion may have implications for joining decisions. We might expect that some treaties would be characterized by an immediate rush to join by a majority of states followed by the gradual acceptance of “hold-outs”. Analytically a logarithmic function form would approximate this process of diffusion. Conversely, diffusion might be


\(^3\) That the joining decisions of other states in a given geographical region positively affect the joining decision of a particular state.
tentative at first followed by an onrush after a threshold point is reached – this form would appear exponential. Another two patterns might emerge between these extremes: diffusion could be linear (that is a constant rate of joining) or logistic (slow early diffusion followed by a rush in the medium term and a slow down in the long term) functional forms. We obtain some improvements in model fit from replacing Simmons’ linear model of diffusion with logistic and logarithmic function forms and provide some evidence demonstrating the applicability of non-linear function forms to other empirical patterns of treaty diffusion.

The article takes the following structure. Firstly, we discuss in more detail the “simultaneity dilemma”. Secondly, we demonstrate the effect of this dilemma on Simmons’ results and then correct for these problems using lags of her covariates. Thirdly, we suggest that using a linear functional form to measure the effect of diffusion may be misleading and we re-analyze the fit of the original model using both logistic and logarithmic versions. Finally, we present empirical evidence that the functional form of diffusion varies across issue areas and suggest that future work on diffusion take this variation into account.

2. The Simultaneity Dilemma

It is an unfortunate but inescapable quandary that not every possible influence on an agent’s behavior lends itself neatly to statistical inquiry. There are various reasons for this limitation. Sometimes this dilemma emerges from reciprocity – that is, we know that the dependent variable is both affected by and also affects the level of the independent variable. Alternatively it might be the case that the level of the dependent variable for one
observation depends on the level of the dependent variable for another observation – that is, there is a violation of the independence assumption. It is particularly unfortunate that measuring diffusion processes can combine both of these problems – reciprocity and dependence – and this predicament derives from the fact that decisions to join treaties are made strategically and concurrently, producing what we refer to as the “simultaneity dilemma”.

The simultaneity dilemma can be modeled as the classic “Stag Hunt” game:

<table>
<thead>
<tr>
<th>Hunt Stag</th>
<th>Catch Rabbit</th>
<th>Join Treaty</th>
<th>Don’t Join</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hunt Stag</td>
<td>5, 5</td>
<td>-5, 2</td>
<td>-5, 2</td>
</tr>
<tr>
<td>Catch Rabbit</td>
<td>2, -5</td>
<td>1, 1</td>
<td>2, -5</td>
</tr>
</tbody>
</table>

In the classic formulation, attributed to Rousseau, a group of hunters are hunting a stag when a rabbit runs across their path. While all the hunters will be best off if the stag is caught and eaten, every person is needed to catch the stag; thus, if one hunter defects to catch the rabbit, the stag will escape and the other hunters will starve. Moreover, the rabbit-catcher still ends up with a more modest meal than his potential share of the stag. This coordination problem is mirrored in the simultaneous decision to join a treaty: if every state in a region decides to join a treaty all will be better off, but if there is a defector it produces negative externalities for all those who chose to join. Hence all players coordinate around not joining or around joining – there is no incrementalism.

If it is the case that simultaneous regional decisions do matter significantly in explaining treaty diffusion this may be extremely difficult to test empirically. The

---

4 In Simmons’ formulation, if all states in a region commits to not placing restrictions on their current account except for one state, this defector could create balance of payments problems for all of the others by reneging on payments due. The dilemma is even more pronounced in the security issue area where the failure of one state to be party to an agreement, for example a ban on weapons of mass destruction (WMD), puts all the other states at risk and prevents coordination around the universally optimal no-WMD solution.
“simultaneity” dilemma emerges because the statistical testing of diffusion requires its operationalization as separate independent and dependent variables. Since explaining diffusion is the focus of our interest, we need to measure each state’s choice to join as a dichotomous dependent variable. However, we also hypothesize that the decisions of other states affect individual states’ decisions to join – thus these other states’ decisions must be modeled as independent variables. The two elements of the simultaneity dilemma elucidated above – reciprocity and dependence – suddenly rear their ugly heads when we attempt this.

The reciprocity problem emerges because in order to create an independent variable measuring the effect of all other joiners at the precise moment a state signs, it is necessary to include those states also signing at that moment. But this means that the dependent variable for some states is incorporated into the independent variable for other states in the same time period. Imagine the decision of state \( i \) is dependent on the simultaneous decisions of all other states \( j \neq i \). To statistically predict the probability of joining at the start of period \( t \) we use the following logit distribution,\(^5\) where \( Y \) is a binary variable measuring the joining decision and \( Z \) is a vector of covariates:

\[
Y_i^t \sim \text{Bernoulli}(\pi_i^t) \quad \text{where} \quad \pi_i^t = \left[ 1 + \exp \left( \beta_0 + \beta_i \left( \sum \limits_{j \neq i} Y_j^t \right) + \gamma Z_i^t \right) \right]^{-1}
\]

Note that the second term in the systematic part of the model incorporates all of the decisions of other states at time \( t \). It is immediately apparent that this allegedly independent variable is actually constructed from the dependent variables for other states.

---

\(^5\) Simmons uses a Cox proportional hazards model so the following demonstration should not be interpreted as an exact formal demonstration of the flaw in her model. Rather the logit distribution is used for ease of explanation.
at time $t$.\textsuperscript{6} This is a serious endogeneity issue – since we are statistically measuring the choice faced by all states at time $t$, we cannot hold constant the simultaneous decision of others to join. In fact our independent variable $\left(\sum Y_{i}^{j}\right)$ is determined by observations of the dependent variable for other states in the same period – therefore we have a reciprocity problem.

The dependence problem also follows from the above formula. We cannot assume that our dependent variable is independently distributed since it is apparent that the covariates used to estimate it depend on the value of other dependent variable observations in the same time period. To simplify, take a case where there are only two states and we are estimating the probability they sign a treaty in time $t$:

\[
Y_{i}^{1} \sim Bernoulli(\pi_{i}) \quad \text{where} \quad \pi_{i}^{1} = \left[1 + \exp\left(-\left(\beta_{0} + \beta_{1}Y_{i}^{2} + \gamma Z_{i}\right)\right)\right]^{-1}
\]

\[
Y_{i}^{2} \sim Bernoulli(\pi_{i}^{2}) \quad \text{where} \quad \pi_{i}^{2} = \left[1 + \exp\left(-\left(\beta_{0} + \beta_{1}Y_{i}^{1} + \gamma Z_{i}\right)\right)\right]^{-1}
\]

It is apparent that if both states sign the treaty, then $Y_{i}^{2} = Y_{i}^{1} = 1$. But according to the Stag Hunt logic outlined above we have no grounds to believe that this is any more likely than $Y_{i}^{2} = Y_{i}^{1} = 0$. Thus we cannot simulate from the model (and derive quantities of interest) since we do not know the value of this “independent” variable. In fact, it is clear that our dependent variables are not independently distributed at all, and thus we cannot use statistical techniques that assume independently and identically distributed data.

Given the simultaneity dilemma, is it feasible to measure diffusion statistically? We argue that it is possible to do so, but it requires the use of \textit{lagged} values of previous levels

\textsuperscript{6} It does not include \textit{previous} decisions made to join at times $t-s$ (where $s=1$): these do not cause endogeneity problems since the dependent variable at time $t$ cannot be affecting things that have already occurred. Thus in this example it makes most sense to assume that previous decisions are incorporated in the covariate matrix $Z$. 
of diffusion rather than attempting to control for simultaneous diffusion, as we show in the next section.

3. Correcting Simmons (2000) for Simultaneity Bias

Simmons (2000) uses a Cox proportional hazards model to examine influences on the decision to sign Article VIII of the IMF Articles, which stipulates that countries should not place restrictions on their current account. Simmons finds a number of substantively interesting results with this model: for example, countries using IMF credits are less likely to join the Article VIII commitments. However, the core hypothesis laid out by Simmons is that “the choices of economic competitors should be important factors” on the choice whether to accede to Article VIII. This claim is operationalized through two independent variables: the “universal” and “regional” proportion of states who have also signed Article VIII. Simmons finds robustly significant results for these two variables: for example, an increase of one percent in the number of all states who have signed is estimated to produce a 5.5% increase in the likelihood of a state signing, and an increase of one percent in the number of states in a given region signing is associated with a 2.7% increase in this likelihood. However, these apparently powerful findings are based on somewhat shaky foundations.

---

7 The current account is the amount owed to other states for goods and services purchased by the home country. Placing restrictions on the current account thus amounts to delaying or reneging on payment for items received. Countries are most likely to replace such restrictions on the current account when they face an incipient balance of payments crisis – that is, inflows on the capital account fail to cover outflows on the current account, necessitating devaluation or bailout.


9 The “universal” diffusion variable measures the percentage of all states who have signed, and the “regional” diffusion variable measures the percentage of states within a region (Simmons uses nine regions) who have signed.
Simmons creates the “universal” and “regional” variables in the following manner: for each country-year, the study records the proportion of states that have acceded to Article VIII, up to and including this year. Recalling our earlier discussion of simultaneity bias it is clear that allowing the value of the “universal” and “regional” diffusion variables to depend partly upon other states signing simultaneously introduces problems of reciprocity and dependence. To account for these time dependencies, Simmons should have used lagged versions of the universal and regional norm variables rather than incorporating other states also signing in a given year. The data used, however, did not include time lags for these variables. Revising the data structure, we have used a one-year time lag for the universal and regional variables, and our results show a substantive difference from those obtained by Simmons.

Before proceeding to replicate Simmons’ Cox proportional hazard tests it pays to examine graphically the changes that lagging the diffusion variables creates in the data structure. Figure One displays density plots of the original ‘universal’ diffusion variable (dashed line) and a lagged version of the same variable (solid line) for all signers. As one can see, the density distributions for the two different versions of the universal diffusion variable are quite different. The version of the data used in Simmons 2000, has a noticeably wider variance than the transformed data set that we used. Simmons’ measure of the “universal” diffusion variable thus displays considerably greater variation on the independent variable than the lagged version, which artificially reduces the

---

10 The density plots show the proportion of states signing (y-axis) at different levels of Simmons’ “universal” diffusion variable and our lagged version (x-axis) for the country-year observation in which a country signed. There are 110 such observations.
standard errors, and inflates the statistical significance, of the obtained coefficient in Simmons, 2000.\footnote{This result obtains because the variance of an estimated coefficient is negatively related to the variance of its respective covariate.}

Table One compares our results to those obtained by Simmons. We find in the first two models that the universal norm maintains a roughly similar statistical significance to that obtained by Simmons\footnote{In fact its magnitude appears slightly stronger although the standard errors are concomitantly larger as we would expect from the smaller variation on the independent variable in the lagged model. See fn. 11. Note, though, that there is a considerable difference in both magnitude and significance for model three.}. Of course, by itself statistical significance is uninteresting. What is does tell us, however, is that the results we obtain using either the lagged or the unlagged versions of the “universal” diffusion variable are extremely unlikely to have resulted by chance. We can be relatively confident that the “universal” diffusion effect is an important determinant of states’ decisions. Thus, the lagged version of the “universal” diffusion variable, although theoretically sounder than Simmons’ original formulation, does not produce any crucial substantive differences.
## Table One – Influences on the Rate of Article VII acceptance

<table>
<thead>
<tr>
<th>Explanatory Variable</th>
<th>Simmons Model 1</th>
<th>Lagged Model 1</th>
<th>Simmons Model 2</th>
<th>Lagged Model 2</th>
<th>Simmons Model 3</th>
<th>Lagged Model 3</th>
<th>Simmons Model 4</th>
<th>Lagged Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universality-lagged</td>
<td>1.066 (0.010)</td>
<td>1.072 (0.019)</td>
<td>1.055 (0.011)</td>
<td>1.060 (0.017)</td>
<td>1.247 (0.089)</td>
<td>1.005 (0.028)</td>
<td>1.040 (0.024)</td>
<td>1.013 (0.020)</td>
</tr>
<tr>
<td>Regional Norm-lagged</td>
<td>1.029 (0.005)</td>
<td>1.0046 (0.004)</td>
<td>1.027 (0.005)</td>
<td>1.0049 (0.005)</td>
<td>1.038 (0.010)</td>
<td>1.0021 (0.007)</td>
<td>1.028 (0.005)</td>
<td>1.010 (0.005)</td>
</tr>
<tr>
<td>Use of Fund Credits</td>
<td>?</td>
<td>?</td>
<td>0.534 (0.160)</td>
<td>0.368 (0.214)</td>
<td>0.577 (0.306)</td>
<td>0.667 (0.169)</td>
<td>0.548 (0.169)</td>
<td>0.380 (0.120)</td>
</tr>
<tr>
<td>Flexible Exchange Rates</td>
<td>?</td>
<td>?</td>
<td>1.52 (0.418)</td>
<td>2.523 (0.700)</td>
<td>2.659 (1.286)</td>
<td>4.103 (1.685)</td>
<td>1.512 (0.482)</td>
<td>1.808 (0.482)</td>
</tr>
<tr>
<td>Surveillance</td>
<td>?</td>
<td>?</td>
<td>0.446 (0.053)</td>
<td>0.458 (0.254)</td>
<td>0.407 (0.295)</td>
<td>0.264 (0.201)</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Openness</td>
<td>1.008 (0.002)</td>
<td>1.0085 (0.003)</td>
<td>1.009 (0.003)</td>
<td>1.0108 (0.0023)</td>
<td>1.019 (0.004)</td>
<td>1.012 (0.003)</td>
<td>1.009 (0.0179)</td>
<td>1.009 (0.003)</td>
</tr>
<tr>
<td>Democracy</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>1.028 (0.034)</td>
<td>1.041 (0.029)</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>GNP/Capita</td>
<td>1.00807 (0.0002)</td>
<td>0.999 (0.000)</td>
<td>1.00007 (0.00003)</td>
<td>0.999 (0.000)</td>
<td>1.00009 (0.0004)</td>
<td>0.999 (0.000)</td>
<td>1.00007 (0.0003)</td>
<td>0.999 (0.000)</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>1.033 (0.020)</td>
<td>1.033 (0.014)</td>
<td>1.035 (0.021)</td>
<td>1.032 (0.015)</td>
<td>1.021 (0.041)</td>
<td>1.041 (0.024)</td>
<td>1.036 (0.022)</td>
<td>1.034 (0.014)</td>
</tr>
<tr>
<td>Reserves/GDP</td>
<td>?</td>
<td>?</td>
<td>1.740 (0.493)</td>
<td>1.833 (0.434)</td>
<td>0.950 (1.192)</td>
<td>0.821 (0.961)</td>
<td>1.744 (0.505)</td>
<td>1.298 (0.315)</td>
</tr>
<tr>
<td>Reserve Volatility</td>
<td>?</td>
<td>?</td>
<td>0.770 (0.157)</td>
<td>0.889 (0.155)</td>
<td>0.883 (0.300)</td>
<td>2.080 (0.736)</td>
<td>0.753 (0.155)</td>
<td>1.027 (0.182)</td>
</tr>
<tr>
<td>Year</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>1.052 (0.051)</td>
<td>1.140 (0.045)</td>
</tr>
</tbody>
</table>

Note: Dependent Variable is dichotomous decision to sign Article VIII of IMF Agreements. Data are Cox proportionate hazard model ratios with time varying covariates. Standard errors are indicated in parentheses. All estimates with p<.05 are in bold. N.B. Cox hazard coefficients are relative to a baseline hazard hence a coefficient of 1.0 is equivalent to no distinguishable effect from the baseline hazard. Coefficients less than one mean a reduced probability of signing, coefficients greater than one mean an increased probability of signing. Probabilities can be directly calculated from the coefficients: hence a one percent increase in universality in Model One Simmons, is associated with a 6.6% increase in the likelihood of signing. Data provided by Beth Simmons (simmons@latte.harvard.edu). N=2651.

However, the models differ from Simmons’ results in one critical way: the effect of the regional norm diminishes in both magnitude and statistical significance. The original effect of regional norms in Model One was measured with a hazard rate coefficient of 1.029, meaning that for every one percentage point increase in the proportion of countries in the same region who join Article VIII of the IMF, the other countries in the region are 2.9 percent more likely to accept Article VIII. The coefficient
that we found measured 1.0046, meaning that the effect was fairly small; for every one percent increase in regional acceptance of Article VIII obligation, other countries which did not accept Article VIII became only 0.5 percent more likely to accede to Article VIII. Moreover, we found much proportionately greater standard errors: in the original model, the robust standard error was approximately one sixth of the measured effect, but in our analysis, the robust standard error is approximately equal in size to the measured effect. Thus, we have no evidence that a robust regional effect is distinguishable. This seems an intriguing finding since it indicates that the regional diffusion effect found by Simmons is more likely to be an artifact of incorrect data analysis than an actual causal effect.

This problem with the estimation of the effect of regional diffusion carries throughout all the models tested. In no case but Model 4 does the coefficient on the “regional” variable become either statistically or substantively significant; in Model 4, the effect of a one percentage point increase in the proportion of neighboring countries leads to a one percentage point increase in the likelihood that non-acceding states will join Article VIII. Even so, the size of the standard error on this coefficient is still proportionately high, equivalent to about one half of the measured effect, and again putting into question the size and direction of the regional norm’s effect. Moreover, this effect only appears once we control for year, which leads to a concomitant decline in the significance of the otherwise robust “universal” variable.

Why is the regional effect so greatly attenuated by the introduction of lags, as compared to the “universal” diffusion effect? One clue can be found by examining the different regional diffusion patterns found across regions. Table Two demonstrates the effect of the simultaneity bias across these various regions.
Table Two – Regional variations in effects of joining Article VIII upon other states

<table>
<thead>
<tr>
<th>Region</th>
<th>Number of States in region who have signed Article VIII by 1998</th>
<th>Average number of other states in region joining when one state commits.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southern Africa</td>
<td>13</td>
<td>1.53</td>
</tr>
<tr>
<td>Western Africa</td>
<td>18</td>
<td>11.67</td>
</tr>
<tr>
<td>East Asia/ Oceania</td>
<td>20</td>
<td>0.2</td>
</tr>
<tr>
<td>South Asia</td>
<td>4</td>
<td>1.5</td>
</tr>
<tr>
<td>Eastern Europe</td>
<td>18</td>
<td>4.11</td>
</tr>
<tr>
<td>Western Europe</td>
<td>21</td>
<td>3.81</td>
</tr>
<tr>
<td>Middle East</td>
<td>10</td>
<td>0.6</td>
</tr>
<tr>
<td>North Africa</td>
<td>3</td>
<td>0.67</td>
</tr>
<tr>
<td>Americas</td>
<td>29</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Averages were calculated by adding the number of other states in a region signing at the same time as each state, then dividing by the total number of states in a given region. A score of zero means that each state in a region joined in a separate year.

Table Two shows the average number of other states from the same region joining Article VIII when each individual state accedes. For example, in Western Africa, when each state joined, on average 11.67 other states joined in the same period, whereas in East Asia/Oceania nearly every state acceded to Article VIII on its own. In Simmons’ dataset the choice of the proverbial 11.67 other states would be included in the “regional” diffusion covariate for a given state in West Africa’s decision to accede to Article VIII. It is unsurprising then that a strong regional effect would be found. However, if we exclude simultaneous decisions by lagging the regional diffusion variable these simultaneous joiners are removed from the covariate and the effect is consequently weakened. If all regions were like East Asia, on the other hand, lagging would make little difference since there are few examples of simultaneous joining. It is the effect of Western Africa, Western Europe and Eastern Europe in particular that biases Simmons’ results toward finding a regional diffusion effect.
Simmons (2000) also examines factors determining the choice of states to place restrictions on their current account, including whether Article VIII accession actually produces any effect on state behavior (the other key finding of the paper and one we do not challenge). She finds, as in the case of joining, that restrictions also follow a regional pattern. Using the same lagging process as earlier we recoded her regional restriction variable so that it did not include the simultaneous choice of other states in the region to restrict but was limited to the proportion of restrictors from the previous period. Table Three demonstrates our results, obtained using a logit model.\(^\text{13}\)

**Table Three – Restrictions on the Current Account**

<table>
<thead>
<tr>
<th>Explanatory Variables</th>
<th>Simmons Model 1</th>
<th>Lagged Model 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-6.99 (.413)</td>
<td>-0.14 (.365)</td>
</tr>
<tr>
<td>Article VIII Commitment</td>
<td>-0.903 (.136)</td>
<td>-1.05 (.116)</td>
</tr>
<tr>
<td>Regional Restrictions</td>
<td>4.00 (.395)</td>
<td>2.35 (.34)</td>
</tr>
<tr>
<td>Terms of Trade Volatility</td>
<td>0.337 (.099)</td>
<td>0.339 (.084)</td>
</tr>
<tr>
<td>Balance of Payments/GDP</td>
<td>-0.016 (.008)</td>
<td>-0.019 (.006)</td>
</tr>
<tr>
<td>GNP per Capita</td>
<td>0.00004 (.00002)</td>
<td>0.00002 (.00002)</td>
</tr>
<tr>
<td>Change in GDP</td>
<td>-0.032 (.013)</td>
<td>-0.033 (.11)</td>
</tr>
<tr>
<td>Openness</td>
<td>-0.002 (.001)</td>
<td>-0.003 (.001)</td>
</tr>
<tr>
<td>Years Since Last Restriction</td>
<td>-1.226 (.108)</td>
<td>-1.236 (.104)</td>
</tr>
</tbody>
</table>

Dependent variable is dichotomous choice to restrict current account in year t. Model is time-series cross-sectional logit. Table replicates Simmons, 2000, Table 4.). Estimates with p<0.05 in **bold**.

Note that most of the coefficients are near identical except for the “regional restrictions” covariate. Because we have lagged the regional restrictions variable it is now no longer taking into account the simultaneous decision of other states in the region to place restrictions on their current account. This is of considerable importance since the

\(^{13}\) It is appropriate to use a simple binary logit model on the whole sample, rather than a case-control rare events model since restrictions on the current account are common, indeed some Article VIII signers like El Salvador have maintained restrictions continuously: see Simmons, 2000. Also King and Zeng, (2001).
placing of restrictions is more analogous to a “Prisoner’s Dilemma” game than a coordination game and thus the stakes are even higher than in the example of treaty accession. However, restrictions can be placed over extended periods of time, thus last year’s regional pattern is a pretty good predictor of the proportion of regional restrictors this year. Balancing these two effects we see a reduced magnitude but continued significance of the regional effect. Figure Two demonstrates 95 percent confidence intervals for the likelihood of a state restricting given various levels of regional restriction for both the unlagged and the lagged version of the covariate. Note that including simultaneous decisions has a much sharper effect on the increase in the probability of a state restricting given a move from low to high levels of regional restriction.

![Figure Two](image)

---

14 In the coordination game (“Stag Hunt”) states that failed to join Article VIII presented a potential threat to the region in that they could choose to restrict their current account. The present example turns on states actually placing these restrictions on their current account. Thus the “sucker’s” payoff is more damaging.
In fact the mean probability of restricting for the unlagged version is lower than for the lagged variable when regional restrictions are less than fifty percent, and is higher when regional restrictions exceed fifty percent. This reflects the Prisoner’s Dilemma logic; given simultaneous decisions, once fifty percent of states are defecting the incentive to also restrict becomes overwhelming – thus the unlagged regression takes a near-logistic form. This threshold effect, however, becomes muted when we look at the lagged variable’s influence, which appears almost linear.

4. Reinterpreting the Pattern of Diffusion

The limitations of Simmons’ regional diffusion variables notwithstanding, the above recoding of the “universal” variable in Table One did not significantly affect the results obtained by Simmons on the influence of cumulative universal diffusion. However, although these results are relatively robust to changes in lags and in model specification, they might be improved by re-evaluating the functional form used by Simmons. The way in which the ‘universal’ variable is coded and entered into regressions assumes a linear effect of diffusion. In other words, a change of the cumulative proportion of signers from zero to ten percent is equivalent to a change from forty to fifty percent, itself equivalent to a change from ninety to one hundred percent. A linear form thus assumes that the cumulative level of signers has a constant effect on non-signatory states. We might ask ourselves whether we find this assumption intuitively attractive. There does not seem to be a good prima facie reason to choose a constant effect over a

---

15 Since this effect holds even controlling for Article VIII membership, it appears that Article VIII is a relatively “soft” form of law with only a moderate effect on state compliance. Regional forces can thus overwhelm the commitment to Article VIII – on “soft” law see Abbot and Snidal (2000).
‘snowballing’ pattern, a ‘petering out’ effect, or some hybrid form.\textsuperscript{16} In fact, we ought to choose the appropriate functional form according to the theoretical assumptions we have made about the process of diffusion.

How ought we to generate hypotheses regarding the process of diffusion? We might presume that some issue areas are more conducive to snowballing (exponential) effects – these would be cases where treaties have strong network effects, that is, they become more attractive the more members they have\textsuperscript{17}. Alternatively one could imagine treaties where most states sign early on but then the rate of acceptance slows as a few stragglers hold out (logarithmic). This kind of treaty, for example that which created the International Criminal Court, tends to be one with heterogeneously distributed costs. One further alternative might be a process with a slow start and slow tail but quick accession in the middle. Such diffusion processes would look S-shaped and thus take a logistic functional form.

We analyzed the rate of acceptance of Article VIII of the IMF to see what kind of diffusion pattern emerged. According to the theoretical logic of Simmons 2000\textsuperscript{18}, we should expect that a “snowballing” effect would emerge because countries join Article VIII in order to provide a credible signal to markets that they provide a good investment environment. Thus, since joining Article VIII is costly we should see an initial reluctance to move followed by a sudden rush to join once competitors choose to accede. In fact,

\textsuperscript{16} A “snowballing” effect starts with initially slow diffusion followed by increasingly rapid acceptance. Conversely a “petering out” effect starts with initially rapid uptake and is followed by increasingly hesitant acceptance.

\textsuperscript{17} Such treaties tend to be those that offer joint goods once a threshold level of members has been passed, at which point the share of these goods outweighs the benefits states get from remaining outside the treaty. A good example from a slightly different field is Congressional declarations of NAFTA support, which started very hesitantly and snowballed as the deadline voting day closed in. See Box-Steffensmeier, Arnold and Zorn (1998).

\textsuperscript{18} In contrast to the empirical operationalization of diffusion in Simmons 2000, which as noted takes a linear form - thus the theoretical expectations clash with their empirical proxies.
when we examine the distribution of the ‘universal’ variable closely, a somewhat different pattern emerges.

Figure Three

Figure Three shows a cumulative distribution function for both Simmons’ version and our lagged version of the ‘universal’ variable. It can be clearly seen that the curve begins with a shallow slope, steepens once about thirty percent of states have signed, and then becomes shallow once more after about sixty percent of states have signed. This logistic shape is not enormously different from a simple linear approximation but it is dissimilar enough to warrant testing a logistic transformation of the lagged universal variable to see if a tighter model fit can be approximated. It should be noted that Simmons’ time-series data did not extend back prior to 1967, although twenty-six states had already joined by that date. Given this lack of observations we also tried a logarithmic transformation (essentially cutting out the first shallow part of the ‘S-shape’) of the remaining data. The results we obtained using logistic and logarithmic functional forms follow in Table
Three\textsuperscript{19}. As can be seen, the fit for the logistic and logarithmic cases is preferable to the linear model fit for the more restricted models but is equivalent or worse for the less restricted models incorporating more covariates.

<table>
<thead>
<tr>
<th>Functional Form</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear (original)</td>
<td>67.89</td>
<td>145.32</td>
<td>74.14</td>
<td>132.38</td>
</tr>
<tr>
<td>Logistic</td>
<td>68.65</td>
<td>145.33</td>
<td>74.19</td>
<td>132.26</td>
</tr>
<tr>
<td>Logarithmic</td>
<td>72.60</td>
<td>142.29</td>
<td>74.11</td>
<td>130.05</td>
</tr>
</tbody>
</table>

Values are log-likelihood ratios calculated for each functional form on the specific model. Higher values indicate better model fit. Tests are Cox hazard proportionate regressions and models are identical to those in Table One save for the replacement of the original linear lagged universal diffusion variable with either logistic or logarithmic forms. All tests are conducted on identical data.

5. The Multiplicity of Diffusion Effects

We have shown that using a logistic and a logarithmic transformation of the lagged ‘universal’ variable for Article VIII data can provide a marginally better fit. However, it was also noted that Simmons’ choice of a linear approximation for the universal variable was not far off. In fact, a linear approximation to the true logistic form of Article VIII acceptance provides a fairly good fit. However other data sets and problems might not work so well. In this section we look at the diffusion patterns for two more treaties to examine whether logistic diffusion patterns tend to be commonplace and how good linear approximations are to real diffusion processes. It is important to note that these are not fully specified or tested hypothetical models; we conduct these tests of the different issue area diffusion rates as a plausibility probe for future analysis of treaty accession rates.

\textsuperscript{19} The logistic form is \((1/1 + \exp[-laguni])^{-1}\) and the logarithmic form is \(\log(laguni)\) where \(laguni\) is the [0,1] lagged proportion of all states previously acceded to Article VIII. N.B. because Simmons’ dataset starts in 1967, the lowest value of \(laguni\) is 0.29 (hence \(\log(laguni)\) never equals -8).
The above figures demonstrate diffusion patterns of diffusion for WIPO and the OTL. N.B. different x-axes are used, the figures are therefore not directly comparable. The WIPO diffusion has the percentage of previous signers on the x-axis, as used above in Figure Three for Article VIII. The OTL treaty conversely had monthly data, which we converted into the exact length of time until signing. The function of these figures thus is to demonstrate the shape of diffusion patterns rather than to allow exact comparison.

The first issue area that we examine also presents as an international political economy matter. The World Intellectual Property Organization (WIPO) is an organ of the United Nations, and it has responsibility for administering 23 different international treaties regarding intellectual property. States that join the WIPO do so with the purpose of gaining transparency in intellectual property matters vis-à-vis other states. As we see in Figure Four (a), the cumulative distribution function for accession to the WIPO is fairly linear. Under modern trade and services regulation treaties, the gains received when one signs onto a treaty are reciprocal – when a state joins, it receives most-favored nation status with all other member states. When a state joins the WIPO or any similar trade treaty organization, the accession does not worsen the state’s interaction with states

---

20 In other words, WIPO provides the kinds of informational services noted by Robert Keohane (1984) that help to resolve states’ otherwise bilateral bargaining dilemmas over property rights.
outside the treaty organization (these remain constant) and it betters interaction with those inside the regime. This differs substantially from the IMF Article VIII case where non-joiners can be damaged once a critical threshold of states acceding to Article VIII has been passed. In the case of the WIPO, however, since there are very few negative externalities to staying out of the organization, states are less likely to be pushed faster than a constant, linear rate by the momentum of others.

In contrast to the areas we have previously examined, which have focused primarily on economic issues, we now examine the Ottawa Treaty (the International Treaty to Ban Landmines or ITBL). Accession to the ITBL has happened in a much more truncated period than the earlier treaties, over months rather than years, and demonstrates a clear logarithmic form. The diffusion curve jumps steeply from the earliest date of accession and continues to rise sharply for the first two years by which time over 50% of the present signers had acceded. After this date the rate slows substantially and is nearly flat after five years. What explains this pattern? As opposed to economic treaties, security treaties present less tangible benefits for states: a landmine treaty is only truly useful if nearly every state has signed. The early signers thus are those for whom signing bears negligible cost (Scandinavian states, Pacific islands), whereas major powers for whom the treaty would bear significant costs have yet to accede, e.g. the US, China, Russia.

6. In Conclusion

What does all the foregoing demonstrate? Several conclusions present themselves. First, and perhaps most importantly, when analyzing diffusion processes, researchers must carefully specify the data format and analytic model to avoid the
“simultaneity dilemma.” Without these necessary changes, which may involve the lagging of diffusion variables, findings on the importance of diffusion effects may not be robust. In particular, we discovered that the regional diffusion result found in Simmons 2000 vanishes once simultaneous decisions are corrected for by the use of lags. While it is certainly possible that diffusion may be lumpy - that is, many states accede at once - such simultaneous decisions cannot be incorporated in statistical analysis because of the joint problems of reciprocity and dependence of observations. This is not merely a problem for survival analysis like treaty-joining but also occurs for more common binary events, for example the choice to place restrictions on the current account. Using a logit model with lags we showed that Simmons overestimated the effect of the proportion of regional restrictors on the choice of individual states to place restrictions.

Finally, we analyzed whether Simmons could have obtained a better model fit in her survival analysis by employing different functional forms for her diffusion variable. We received mixed results with this change: more restricted models tended to produce a better model fit when nonlinear diffusion forms were used, however, in less restricted models the fit of nonlinear variables was roughly equivalent to the linear version. By briefly examining the diffusion patterns of two other treaties (the WIPO, and the ITBL) we show further examples of variation in diffusion patterns and suggest that future statistical analysis of treaty accession take into account these patterns when measuring the effects of past diffusion. This calls into question the assumption that joining occurs at a constant rate and should provoke analysis of the importance tipping points and momentum in the analysis of the spread of international institutions and agreements.
References


King, Gary and Langche Zeng, ‘Explaining Rare Events in International Relations Data’, *International Organization*, 2000, 54.


Simmons, Beth A., ‘International Law and State Behavior: Commitment and Compliance in International Monetary Affairs’, *American Political Science Review*, 2000(b), 4