Can the number of veto players measure policy stability?

Monika Nalepa and Ji Xue (The University of Chicago)

March 16, 2019

Abstract

Ever since the publication of George Tsebelis's "Veto Players", political scientists have constructed measures of policy stability on the basis of his theory's implications. In this research note, we indicate two popular misunderstandings associated with this use. To some extent these problems are known, but they scholars persist in ignoring them nevertheless, warranting drawing our attention to them. The first is the failure to distinguish between weak and strong implications of veto player theory. The second misunderstanding lies in the fact that scholars ignore information about the preferences of veto players in different settings. Focusing on the sheer number of veto players is particularly dangerous when constructing measures that are applied to draw inferences from cross-sectional comparisons. After explaining the nature of these misunderstanding using a stylized example, we provide evidence from two popular datasets proposing measures of policy stability that have been used in political science journal publications. We also develop a set of best practices for using the number of veto players as a measure of policy stability.

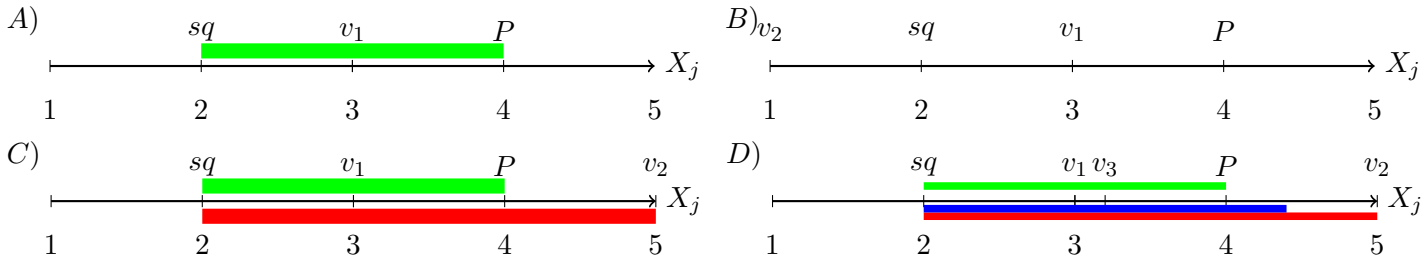
1 Introduction and example

According to George Tsebelis (2002), every political system can be characterized by the number, preferences, and proposal power of its veto players. Consider, for simplicity, a uni-dimensional policy space. Given a status quo policy x and a set of veto players characterized by ideal points in this policy space, one can define a *winset*, $W(x)$, which is the intersection of the sets of policies that each veto player would prefer to x. One can show that in order for a policy to be implemented over x, it must belong to $W(x)$. Because adding veto players can only shrink, but cannot expand, the winset, Tsebelis predicts that adding veto players to a political system will never decrease but can only increase or leave unchanged policy stability.

This allows for a more general and useful classification of political regimes than into parliamentary, presidential, and semi-presidential. Unfortunately, most political scientists who apply this theory to measure political stability focus exclusively on the number of veto players, ignoring the other two characteristics of proposal power and preferences. This leads to inappropriate uses of the measure, particularly in cross-sectional settings. This article discusses the consequences of this omission using as examples the measures from two influential datasets:: *The Database of Political Institutions* (Beck, Clarke, Groff, Keefer & Walsh 2001) (CHECKS) and the *Political Constraints Index* (Henisz 2002) (POLCON).

We use a stylized example to motivate our argument. In all our graphs, the winset is the overlapping region of all the colored lines in the graph. Consider part A of figure 1 below, showing the ideal point of a single veto player, v_1, alongside exogenously given status quo, sq, and proposal, P. At the bottom of each panel, we have included cut points at regular intervals from 1 to 5 for ease of gauging spatial distance. If v_1 is the only veto player in this setting, proposal P, equidistant from v_1 and sq, will be accepted (as will any proposal between sq and P). v_1,acceptance region is marked by the green line. Consider next the addition of another veto player, whose ideal point is labeled with v_2, as represented in panel B of figure 1. Now proposal P would get rejected. In fact any alternative proposal to sq would be rejected. The addition of veto player v_2 has increased policy stability in line with Tsebelis's theory. Suppose, however, that veto player v_2 had been added not at position "1" but at position "5", as illustrated in example C. Now, the addition of veto player v_2 has no effect on policy stability relative to example A. It is still the case that any proposal between

Figure 1: How veto players affect policy stability: stylized example



sq and P will be accepted by these two veto players, as is shown by the overlapping region of green line and red line. Thus, whether or not adding veto players actually affects policy stability depends on the preferences of veto players relative to the status quo and to other existing veto players in the system.

Another way to see this is to compare example C with example D. In C, there are just two veto players: v_1 and v_2. Yet, C has just as much policy stability as does the system in example D, where there are three veto players. (In D, the winset is the overlapping region of the red line, green line, abd blue line.)This aligns with Tsebelis's theory that the addition of veto players will never decrease policy stability. However, many empirical works in political science cite the veto player theory when backing expectations that increasing the number of veto players ought to *strictly increase* policy stability. For the purpose of this note, we refer to this error as *misunderstanding I*.

Note, further, that when Tsebelis talks about the addition or reduction of veto players, he has in mind the addition or subtraction of veto players only relative to the same structure. His theory does not imply that claims about policy stability can be made through the comparison of numbers of veto players in different preference structures. That systems with more veto players do not necessarily have more policy stability is illustrated by comparing example B to example D. While D has three veto players and B has two, D has less policy stability than B. We will refer to this error of drawing inferences on policy stability from the cross-sectional comparison of the numbers of veto players alone as *misunderstanding II*. Admittedly, *misunderstandings I* and *II* are well-known among political science scholars, but they are nevertheless committed with alarming frequency in empirical research simply because solutions to these problems have not been developed.

2 Use of Veto Points Theory in Empirical Political Science

15 years ago, Steffen Ganghof (2003) anticipated difficulties with the application of Tsebelis's theory, identifying three problems: the failure to identify the true veto players, to measure these players' preferences, and to identify different types of veto players. The first problem is associated with focusing on institutional veto players without accounting for informal veto players. As an example, consider a cabinet in a parliamentary regime that is a coalition made up of three parties. A measure of policy stability would need to count each member of the ruling coalition as a veto player. Furthermore, if one of the parties making up the ruling coalition were to be divided into factions, the measure would also have to reflect the number of these factions. The second problem boils down to what we identify above as misunderstanding I. Failure to account for veto player preference may lead scholars to predict a reduction in stability where Tsebelis's theory, noting that the additional veto player whose ideal point falls in between veto players with extreme ideal points does not impact the range of proposals preferred to the status quo, predicts no change. As an illustration, compare a pair of transitions in figure 1. The transition from A to B is a veto player addition that decreases policy stability, while the transition from C to D is a veto player addition that has no effect on policy stability. Failure to account for the difference between preferences represented by v_3 from D relative to C and those represented by v_2 from B relative to A constitutes misunderstanding I, or the second pitfall Ganghof identifies.

The third problem is arguably a problem with Tsebelis' theory itself, which does not distinguish between types of veto players. A constitutional court has the same status as does a leader of a party faction, even if the latter is completely informal and even if his veto intervention capacity can only be mobilized under special circumstances. Because both of the datasets (that use veto players to measure policy stability) discussed here include other variables allowing scholars to account for regime type, we do not elaborate on Ganghof's third error.

Despite Ganghof's early and significant contributions, many scholars have continued to measure policy stability using the number of veto players.

Yet, even if scholars correctly identify all political actors with formal or informal veto powers, adequately represent their preferences (so avoiding misunderstanding I), and control for regime type, they may still engage in misunderstanding II, a mistake even Ganghof failed to acknowledge.

Avoiding all of the pitfalls Ganghof lists does not shield scholars from making unwarranted predictions based on cross-sectional comparisons of veto player number. Whether or not scholars use readily-available veto player data correctly or incorrectly depends on the specific statistical method they choose.

The remainder of this note will focus on the use of two measures from datasets popular in IPE and CPE. We identify how many articles use these measures correctly, and present data on how often the use of these measures entails mistakes I and II. We also outline a way forward for the correct use of measures of policy stability using the concept of veto players.

2.1 CHECKS1 as a measure of policy stability

Among three measures proposed by Beck et al. (2001) we focus on CHECK1, the most commonly used in the political economy literature. The variable counts the number of veto players in a political system correcting for preference alignment determined by three factors: (1) the number of parties competing in elections; (2) the electoral system; and (3) the party affiliation of the veto players. These adjustments are made by mechanically adding one to or subtracting one from the number of veto players. Scholars who use this measure are likely to commit both misunderstanding I and misunderstanding II. First, using this measure almost inevitably leads to misunderstanding II, because its construction ignores ideological positions. Second, it wrongly predicts a strict increase in policy stability when more veto players are added, while Tsebelis's theory only predicts that stability will not decrease with an increase in the number of veto players unless preferences of veto players are accounted for. Therefore, studies where CHECKS1 appears insignificantly different from zero may not be refuting veto players theory, as this theory indeed may predict that adding veto players has no effect.

Misunderstanding I will be committed whenever such an inference is drawn regardless of whether scholars are making cross-sectional *or* time series comparisons;

For example, suppose a scholar were to use CHECKS1 in a time series regression to show that increasing the number of veto players makes land reform less likely to occur in a given year. Further, suppose this scholar were to find that CHECKS1 has no significant effect in predicting the probability of land reform. Were this scholar to claim their finding to be inconsistent with veto players' theory, they would be committing misunderstanding I, as a lack of change in policy sta-

bility with the addition of a veto player is consistent with veto players' theory.

Furthermore, since CHECKS1 does not account for the distribution of ideological positions, scholars who use the CHECKS1 measure will almost certainly be committing misunderstanding II.

2.2 The Political Constraints Index

The second prominent dataset proposing a variable based on the number of veto players to measure political stability is the Political Constraints Index (Henisz 2002). In order to construct the variable (*POLCON*), Henisz first extracts from political science databases the number of independent branches of government (focusing on the executive and lower and upper legislative chambers, as these presumably have veto power over policy). Next, he assumes that the preferences of each of these branches are independently and identically distributed over a unidimensional policy space. For instance, for $n = 2$ veto players, this distance is $\frac{1}{4}$ because $\frac{1}{(n+2)} = \frac{1}{4}$ (Henisz 2002, 381). The assumption that the preferences of each veto player are IID is strong, especially in parliamentary regimes, where the executive is a direct outcome of legislative elections and must be aligned, at least to some extent, with the preferences of the legislative median party. Aware of this, Henisz adjusts the measure to take into account the partisan affiliation of the executive, the size of the legislative majority, and the fragmentation of the legislature (Golder 2006). Since POLCON accounts for directionality (the side of the existing veto players to which additional player falls), the addition of veto players within the same country over time is measured well. As a result, it should work well for time series analysis within the same country and misunderstanding II is avoided as long as the model includes cross-section "country" fixed effects.

In other words, it is warranted to use this measure to claim that: "Countries are more likely to implement land reform as they become more authoritarian than as they become more democratic" (Albertus 2011), where becoming more authoritarian and more democratic are operationalized by a decrease and increase in the number of players (as measured by POLCON), respectively. It is not, however, warranted to claim that "countries with more veto players are more likely to implement land reform than countries with fewer veto players."

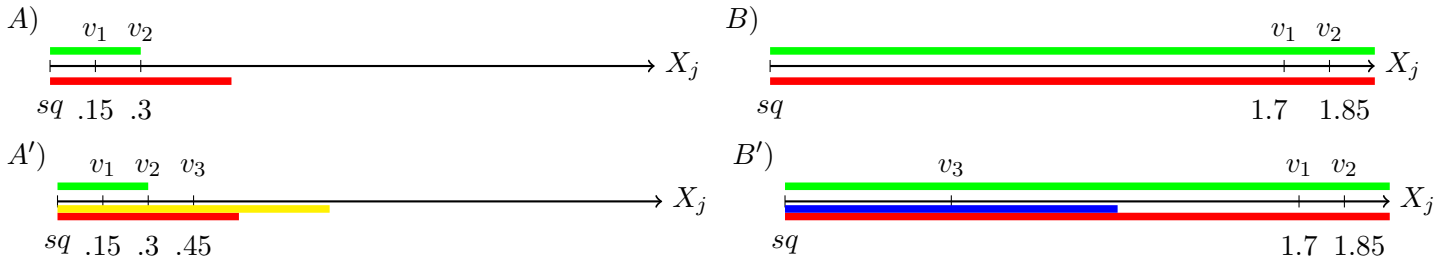
An illustration of this common mistake from political science is provided by Mansfield, Milner & Pevehouse (2007). The authors of this article analyze country dyads who have the opportunity

of entering a Preferential Trade Agreement (PTA), and find that the probability of a PTA forming declines with the number of veto players in each country. Consider three countries, corresponding to examples A, B, and D from Figure 1, and assume that P represents the spacial location of the PTA agreement. Veto players' theory predicts that it would be easier to form a PTA between countries A and D than between countries A and B, but D has 3 veto players while B has only 2. Hence, holding constant the number of veto players in country A, the country with *more* veto players ought to be entering a PTA with higher probability than the country with *fewer* veto players, the opposite of what (Mansfield, Milner & Pevehouse 2007) expect to find.

We use another example, illustrated in figure 2, to show why the measure cannot be used cross-sectionally. The distance between veto players in example A and in example B is .15; from the point of view of POLCON, then, their structures are identical. If the status quo is to the left of both veto players (for instance, at 0, where sq is located at 0 in A and B), the addition of veto players in example A is much more likely to occur to the right of both veto players(shown in A'), which is unlikely to affect policy stability (note that the intersection of winsets will be the interval $[0, 1.5]$ regardless of the inclusion of additional veto players). In A', the overlapping region is the same as in A. However, in example B, the addition of veto players(shown in B') is much more likely to occur to the left of v_1 and v_2 and be consequential for stability: every addition to the left of the left-most veto player in example B will further shrink the winset, which is the overlapping area of all the colored lines in the figure. Compared to B, in B' the winset shrinks. This is why the two structures, even though treated as identical by Henisz's measure, are not the same. In a nutshell, Henisz's measure accounts for relative distances between existing structures of veto players, but does not take into account their moderation or extremity. This is particularly relevant for uses of POLCON in comparative politics, where regime types ranging from inclusive democracy to harsh dictatorship are plausible interpretations of such divergent structures as shown in figure 2.

Before providing aggregate evidence for the misuses of veto player-based measures of policy stability, we mention two other particularities of the Political Constraints Index. First, it ignores the judiciary as a distinct type of veto player. This is particularly significant for the work of comparativists. Stalling decisions made by constitutional courts are just as consequential for policy stability as a presidential veto or the failure of an upper chamber to corroborate a lower chamber decision. However, in many contexts, constitutional courts cannot veto legislation unless explicitly

Figure 2: The assumption of the uniformly distributed status quo



requested to do so by a political actor and must justify any vetoes with constitutional provisions. Ignoring the specificities of courts decreases the reliability of the Political Constraints Index. A second problem is the failure to distinguish between presidential and parliamentary regimes. Even if the president is from the same party as a majority of the legislature, their separation of origin and survival can make them uneasy bedfellows (Samuels & Shugart 2010). For anecdotal evidence one need look no further than the struggles of Donald Trump to muster a Republican majority for repealing Obamacare. A prime minister leading the party holding a legislative majority no doubt has an easier time getting the support of his party's members in the legislature. In order to remedy some of these problems, Henisz developed , POLCON V, which follows the same logic as POLCON III but accounts for two additional veto players: the judiciary and sub-federal entities. Though this measure considers potential veto players more thoroughly, it should also not be used for cross-sectional comparisons.

It should be noted that Henisz does not, in his article presenting the dataset, encourage scholars to apply it cross-sectionally; indeed, as previously mentioned, he takes care to account for veto player preferences, recreating the party composition of each legislature and adjusting alignment with the fractionalization index to reflect how far apart veto players are from one another. Yet, even leaving these nuances aside, the inconvenient conclusion of the paragraphs above is that although POLCON may be used for assessing the relative stability of the same unit over time, it cannot be used for making cross-sectional inferences.

This does not mean that any scholar working with panel data is automatically "off the hook" for committing misunderstanding II. In order to ensure that the effect of veto players is estimated separately for each country, scholars have to include country fixed effects in their regressions. Implementing random effects or panel corrected standard errors takes care of heteroscedasticity is-

sues but does not ensure the correct use of the number of veto players. The penultimate section of this note presents the results of a meta-analysis of recently published political science articles citing either of the two databases discussed above in a use of veto player numbers in their empirical analysis, and classifies them into (1) papers that use the CHECKS1 measure and commit misunderstanding I, (2) papers that use CHECKS1 measure and do not commit misunderstanding I, (3) papers that use the POLCON measure incorrectly and commit misunderstanding II, and (4) papers that use the POLCON measure correctly.

3 Meta-analysis of political science articles

Table A1 in the Appendix lists papers that used the CHECKS1 measure.[1] To see whether a paper engaged in misunderstanding I, we verify if the following 3 conditions are met: (1) CHECKS1 is used as a proxy for policy stability/flexibility or political constraints, (2) CHECKS1 turns out to be insignificant in empirical analysis, and (3) the author explicitly or implicitly points out that the insignificance of CHECKS1 means that veto players/policy flexibility/political constraints have no effect on the outcome variable of interest. All three conditions must be satisfied in order for misunderstanding I to occur.

Since the construction of the CHECKS measure does not account for veto player preferences, all the articles that used CHECKS as a measure for political constraints are liable of misunderstanding I and II. *All* of them, but one (Crespo-Tenorio, Jensen & Rosas (2014)) commit this error.[2].

Table A2 focuses on misunderstanding II, listing all articles using POLCON V(occasionally III) that meet our selection criteria(published after 2000; having at least 10 citations if published before 2015). Column "DV scale" indicates whether the dependent variable is dichotomous, continuous, categorical, or treated "as if continuous." A DV is categorized "as if continuous", if although discreet in nature, it can take up many values. Among articles listed in Table 2, 22(almost half) are guilty of misunderstanding II while 24 avoid it. Those engaged in misunderstanding II typically either use a dichotomous or categorical dependent variable. Misunderstanding II was unavoid-

[1]More details on the search results are provided in the Appendix.

[2]Crespo-Tenorio, Jensen & Rosas (2014) uses CHECKS to measure clarity of responsibility and not policy stability

able in articles that, due to data limitations (too short a time series), had to rely on cross-sectional comparisons.

No papers in our sample relied on a time series alone, though ten relied exclusively on cross-sectional comparisons. These papers, for obvious reasons, could not include fixed country effects. Yet, even though they had the opportunity to avoid it, many of the POLCON papers that analyzed panel data subjected themselves to misunderstanding II by refraining from fixed effects. Why? The avoidance of fixed effect analyses is most likely attributable to one of two reasons. First, if the dependent variable is not continuous and shows no variation in some countries, fixed effects would lead to throwing those countries out of the panel. Thus, papers that use dichotomous and categorical DVs may be especially prone to avoiding fixed effects. Articles with continuous DVs may face a similar problem if their DVs vary little within the same countries. Second, and relatedly, panels with short time series but wide cross-sections will lose relatively many degrees of freedom when including a dummy for every country. Again, scholars worried about losing degrees of freedom may settle on random effects or panel corrected standard errors, both of which make them vulnerable to misunderstanding II. Since these considerations are important, we decided to include in Table A2 information about the DV as well as the length of the time series.[3]

4 Preliminary Examinations of Current Stability Measures

In this section, we did some simple, preliminary examination on how the current policy stability measures perform against legislative output data which shows how often policy changes have taken place.

First, We can see how badly CHECKS perform using an extremely simple and crude legislative output data: how many bills passed and became law each year in one single country. The data we used is the number of UK Public General Acts (bills that passed) each year from 1975-2016[4], and POLCON V and CHECKS1 index of U.K. in the same period.

To form a comparison, POLCON V is supposed to be non-problematic when it applies to time-

[3]Detailed explanations of why a particular paper was classified in a certain way have been relegated to the electronic appendix.

[4]Data is from U.K. parliament website. https://www.legislation.gov.uk/ukpga

Table 1: Acts, Checks, and POLCON Data

year	checks1	POLCON	acts
1975	4	0.740771111294419	65
1976	4	0.740771111294419	62
1977	5	0.740771111294419	32
1978	5	0.740771111294419	44
1979	5	0.740771111294419	48
1980	4	0.738317143595221	53
1981	4	0.738317143595221	60
1982	4	0.738317143595221	46
1983	4	0.738317143595221	50
1984	4	0.736438685551736	53
1985	4	0.736438685551736	68
1986	4	0.736438685551736	62
1987	4	0.736438685551736	49
1988	4	0.738762360222828	55
1989	4	0.738762360222828	46
1990	4	0.738762360222828	46
1991	4	0.738762360222828	69
1992	4	0.738762360222828	61
1993	4	0.741330140186695	52
1994	4	0.741330140186695	41
1995	4	0.741330140186695	54
1996	4	0.741330140186695	63
1997	4	0.741330140186695	69
1998	3	0.737176826913764	49
1999	3	0.737176826913764	35
2000	3	0.737176826913764	45
2001	3	0.737176826913764	25
2002	3	0.738442413715171	44
2003	3	0.738442413715171	45
2004	3	0.738442413715171	38
2005	3	0.738439338917306	24
2006	3	0.746082340317277	55
2007	3	0.746082340317277	31
2008	3	0.746081060359517	33
2009	3	0.746078500443996	27
2010	3	0.746074660570715	41
2011	4	0.748529252862392	25
2012	4	0.748529252862392	23
2013	NA	0.748529252862392	33
2014	NA	0.748529252862392	30
2015	NA	0.748529252862392	37
2016	NA	0.748529252862392	25

series data, which is what we observe here. There is a negative correlation of -0.51 between POLCON V and number of bills passed each year. And the scatterplot of Polcon V and the number of bills passed (Figure 3) shows a clear downward trend: fewer bills passed, more stability as predicted by POLCON, as is predicted.

1 acts and polcon5.png

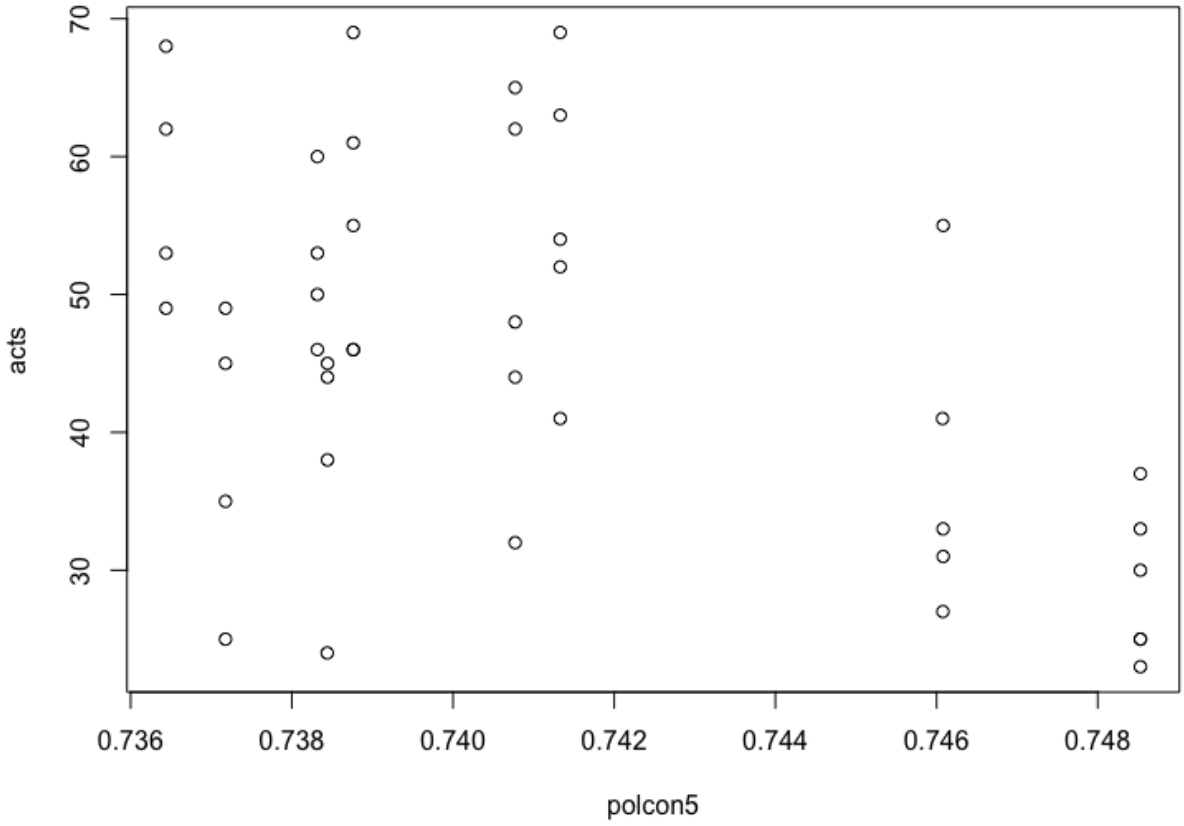


Figure 3: scatterplot of POLCON and the number of bills passed (General Public Acts) each year in U.K. 1975-2017.

By comparison, Checks1 performs rather poorly with our legislative output data, indicating that it is perhaps not a very good measure of policy stability[5]. The correlation between the number of acts and checks is 0.346, which is completely off because a higher Checks1 index is supposed to lead to fewer acts. In the scatterplot, if checks1 performs well, we are supposed to be able to

[5]CHECKS is updated till 2012, so we used data from 1975-2012.

see a strong downward trend similar to the POLCON plot. On the contrary, instead of the clear downward trend of the POLCON scatterplot, we can barely see any correlation at all (Figure 4).

2 acts and Checks1.png

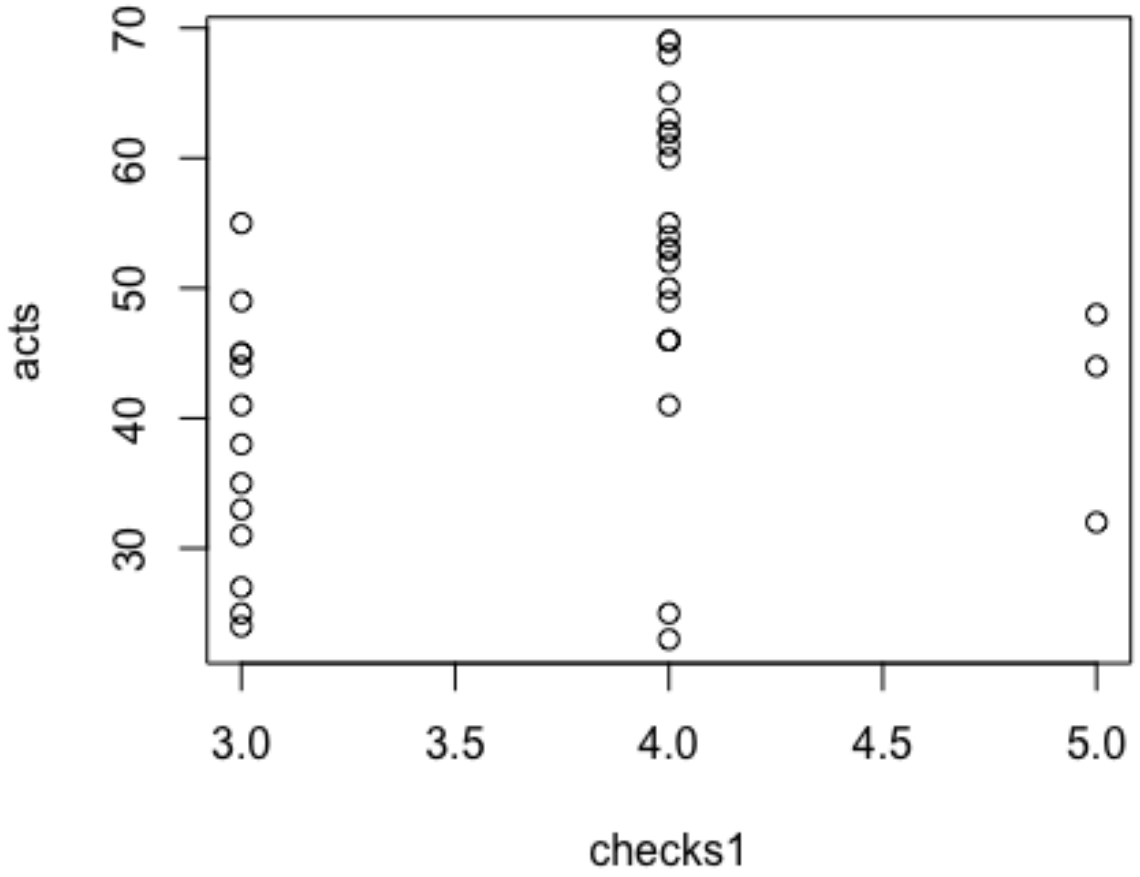


Figure 4: acts and checks1

But even though POLCON does well with time-series data, this index performs terribly when it is used for pooled data of more than one country. POLCON is never supposed to be used cross-nationally. To see that, we constructed a pooled dataset of 13 Western European countries for a period of 20 years. We did not use the number of bills passed each year, because some countries tend to pass considerably more bills on average than other countries, even when these countries have

the same type of political system. Therefore, we used Strobl's data on reform, which gave us the number of significant policy reforms carried out each year in 13 Western European countries from 1985-2005[6]. Only important reform measures are included. This dataset includes 5600 reforms, covering areas of social, taxation, labor and economic policy. The countries are Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Portugal, Spain, Sweden and the United Kingdom. In this pooled dataset, there are 292 observations. The number of reforms each year ranges from 1 to 52 (see Table 2).

Table 2: Summary of Reform and POLCON Dataset

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
countrycode	292	7.459	4.199	1	4	11	14
year	292	1,995.281	6.819	1,979	1,990	2,001	2,009
reforms	292	17.716	9.194	1	11	23	52
polcon	292	0.789	0.063	0.341	0.755	0.843	0.894

If POLCON performs well, we should see a strong negative correlation between the two variables: as number of reforms gets higher, POLCON is supposed to get lower. However, this is not the case: the correlation between POLCON V and reforms per year is a poor -0.040318. The scatterplot of these two variables (Figure 5) does not show any pattern at all. When we use POLCON to predict reforms in a bi-variate regression, POLCON is not significant (see Table 3).

5 Ways forward

These results indicate that the incorrect use of veto players as a measure of policy stability is relatively common in empirical political science work. These misunderstandings could, however, be avoided. As a rule, scholars using panel data should always aspire to including fixed effects in their regressions to ensure that countries with different veto player structures are not being compared to one another.

[6]The reform data is from the following paper: Angelova, Mariyana, et al. "Veto player theory and reform making in Western Europe." European journal of political research 57.2 (2018): 282-307. The dataset can be downloaded here: https://homepage.univie.ac.at/daniel.strobl/replication-data.html

3 reforms and POLCON.png

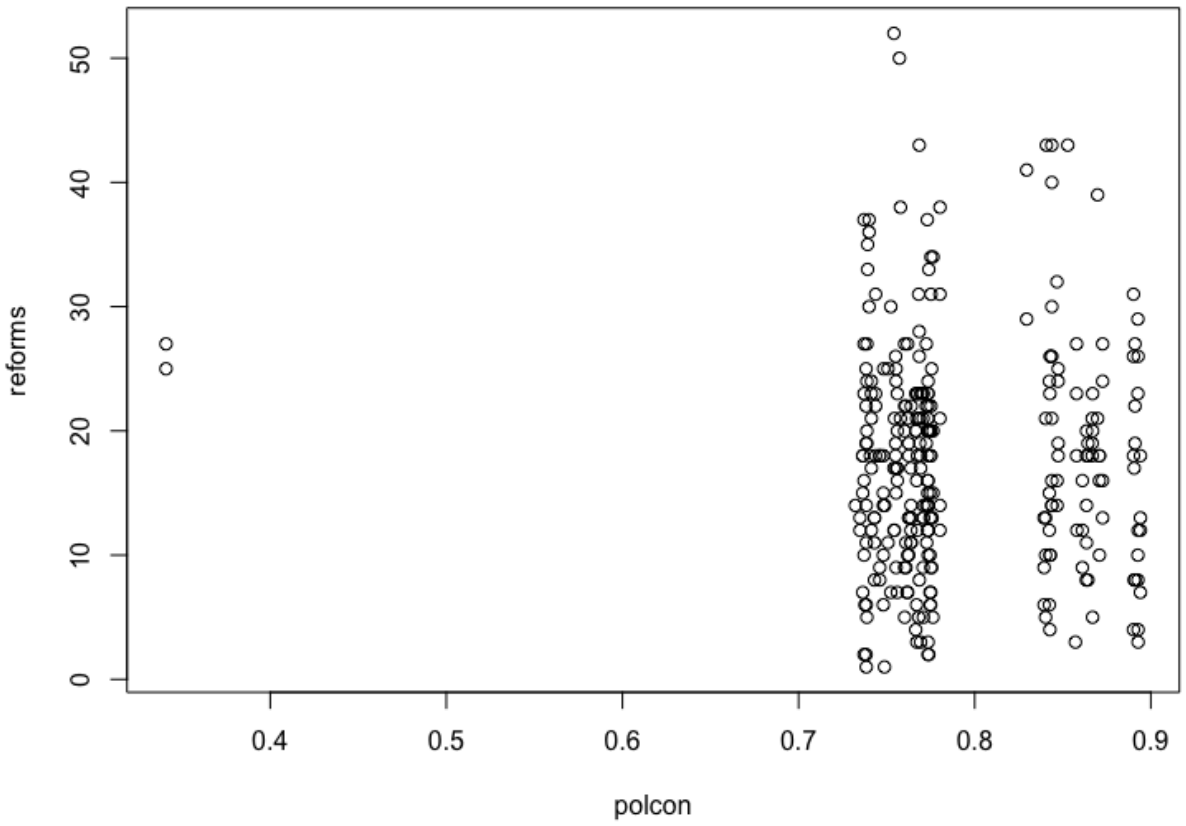


Figure 5: reforms and POLCON

In some circumstances, adding fixed effect is not feasible, which occurs when the length of a time series is too short to warrant such an operation, or if there is insufficient variation in the dependent variable in countries scholars do not want to throw out of the panel. Occasionally, if the time series for a country is very long, significant institutional changes within a country can modify both the number and identity of VPs. Then, even a single country time-series can fall prey to Misunderstanding II. The latter situation is rare but possible. In these circumstances, scholars must look beyond veto player number to measure policy stability.

A research note format does not permit us the space tp fully develop as solution to this conundrum, but we sketch two potential ideas for such solutions.

The first relies on calculating to what extent potential veto players agree or disagree with one

Table 3: Reforms and polcon

	Dependent variable:
	reforms
polcon	−5.929
	(8.629)
Constant	22.392***
	(6.826)
Observations	292
R^2	0.002
Adjusted R^2	−0.002
Residual Std. Error	9.203 (df = 290)
F Statistic	0.472 (df = 1; 290)
Note:	*p<0.1; **p<0.05; ***p<0.01

another. On the one hand, no matter how many potential veto players there are, if a set of players rarely disagree with each other when they vote, the policy flexibility is probably high and changes can be easily made. On the other hand, if a set of players often voice very diverse opinions and are constantly in conflict with each other during voting, then we could foresee policy stability. Modern techniques of ideal point estimation based on Item Response Theory can be applied to any roll-call data at the level of assemblies and scholars have been successful at applying it to reconstruct preferences of legislators (Poole & Rosenthal (2000); Clinton, Jackman & Rivers (2004)), judges (Martin & Quinn (2002);Hanretty (2012)) and even UN ambassadors (Mattes, Leeds & Carroll (2015);Hug & Lukács (2014)). Ideal point estimation techniques, though designed to apply to roll-call voting, could also be applied to institutions that are equipped to stop or move forward any proposal along the legislative process. In other words, actions of lower and upper assemblies, presidents, and courts could be treated as ideal point estimation treats legislators voting in assemblies. This technique, when applied to a common set of proposals—such as international treaties or trade agreements—would allow scholars to compare the actual amount of disagreement in each country and at least produce an ordering over the countries in question with regard to the potential policy stability of countries.[7]

[7]There are potential problems with such measures. Although seemingly exogenous, proposals

A second way forward would be to use ideal point estimation on traditional assembly data in each country, but only to find the location of the players with the most extreme preferences vis-a-vis each other. These locations could be used to define the Pareto set for the assembly, in which only these most extreme players are relevant. The size of Pareto sets would serve as a policy stability baseline for each country. The next step would require the use of historical data over time, to see how the size of this Pareto set changes over time. If it does not change much, it is probably because it already started at a level of high stability. If it is sensitive to change, the policy stability was likely low to begin with. As with the previous measure, this operation would produce an ordering over countries from the most stable to the least stable. (For a simple proof of why Paretoset can be a good measure of policy stability, see the Appendix: "Size of Pareto set as measure of policy stability").

A potential obstacle to using this measure is that the Pareto set is hard to define in a multi-dimensional setting. But we have reason to believe multiple dimensions are rare because of the way attitudes across issues are constrained: In assemblies, the number of dimensions needed to represent ideal points is typically small, because legislators decide how to vote on the basis of their positions on a small number of underlying evaluative, or basic, dimensions. For example, in recent U.S. congresses, we can easily predict how a "liberal" or a "conservative" will vote on most issues. So in the U.S. setting, veto players' positions are mostly uni-dimensional. [8] Therefore, by looking at voting records, we can easily identify the two things we need in order to construct such new measures, based on who are the most extreme players, and how far apart are their positions, and how these distance change over time.

---

of international treaties and trade agreements may favor certain countries over others, leading them to agree more readily than others.

[8]Poole (2005), in his classic book on spatial theory of ideology, argues that, if voters have coherent belief systems, their positions on a variety of specific issues can be captured by her position on one or two fundamental dimensions such as liberalism-conservatosm.

References

Albertus, Michael. 2011. "Democracy and the threat of redistribution in Latin America.".

Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer & Patrick Walsh. 2001. "New tools in comparative political economy: The database of political institutions." *the world bank economic review* 15(1):165–176.

Clinton, Joshua, Simon Jackman & Douglas Rivers. 2004. "The statistical analysis of roll call data." *American Political Science Review* 98(2):355–370.

Crespo-Tenorio, Adriana, Nathan M Jensen & Guillermo Rosas. 2014. "Political liabilities: surviving banking crises." *Comparative Political Studies* 47(7):1047–1074.

Ganghof, Steffen. 2003. "Promises and pitfalls of veto player analysis." *Swiss Political Science Review* 9(2):1–25.

Golder, Matt. 2006. "Presidential coattails and legislative fragmentation." *American Journal of Political Science* 50(1):34–48.

Hanretty, Chris. 2012. "Dissent in Iberia: The ideal points of justices on the Spanish and Portuguese Constitutional Tribunals." *European Journal of Political Research* 51(5):671–692.

Henisz, Witold J. 2002. "The institutional environment for infrastructure investment." *Industrial and corporate change* 11(2):355–389.

Hug, Simon & Richard Lukács. 2014. "Preferences or blocs? Voting in the United Nations human rights council." *The review of international organizations* 9(1):83–106.

Mansfield, Edward D, Helen V Milner & Jon C Pevehouse. 2007. "Vetoing co-operation: The impact of veto players on preferential trading arrangements." *British Journal of Political Science* 37(3):403–432.

Martin, Andrew D & Kevin M Quinn. 2002. "Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis* 10(2):134–153.

Mattes, Michaela, Brett Ashley Leeds & Royce Carroll. 2015. "Leadership turnover and foreign policy change: Societal interests, domestic institutions, and voting in the United Nations." *International Studies Quarterly* 59(2):280–290.

Poole, Keith T. 2005. *Spatial models of parliamentary voting*. Cambridge University Press.

Poole, Keith T & Howard Rosenthal. 2000. *Congress: A political-economic history of roll call voting*. Oxford University Press on Demand.

Samuels, David J & Matthew S Shugart. 2010. *Presidents, parties, and prime ministers: How the separation of powers affects party organization and behavior*. Cambridge University Press.

Tsebelis, George. 2002. *Veto players: How political institutions work*. Princeton University Press.