European Union Politics

http://eup.sagepub.com/

Assessing the Contributions of the DOSEI Project

George Tsebelis

European Union Politics 2005 6: 377
DOI: 10.1177/1465116505054838

The online version of this article can be found at:
http://eup.sagepub.com/content/6/3/377

Published by:
SAGE
http://www.sagepublications.com

European Union Politics
DOI: 10.1177/1465116505054838
Volume 6 (3): 377–390

SAGE Publications
London, Thousand Oaks CA,
New Delhi

Forum Section

Assessing the Contributions of the DOSEI Project

George Tsebelis
University of California, Los Angeles, USA

The DOSEI (Domestic Structures and European Integration) group was faced with one of the most important real-scale political realities of our time: the creation of the constitution of today's second-biggest economic power. In scope and significance there is one parallel in history: the creation of the American Constitution in 1776. Understanding how the EU Constitution emerged is one of the major tasks of political science.

Although the significance of the political events to be explained is overwhelming, a number of challenging methodological issues required solutions before answers to the equally challenging political questions could be tackled. The DOSEI researchers had to identify the preferences of the actors involved. This is a difficult task because, by definition, preferences are features of individual actors, *known* only to them, and they have to be *inferred* by all other actors and observers.

Why describe the preferences of these actors in terms of 'inference' instead of 'knowledge'? Why can we not ask actors what they prefer, and then use their answers to analyze their actions? There are obvious technical problems with such a procedure: the actors themselves are busy with other tasks, and experts may get wrong impressions. Moreover, the answers may be incomplete, particularly for actors no longer on the scene, or researchers may not have been interested enough to collect the information. Besides, a serious analysis has to confront one fundamental impediment: the interaction with *strategic* actors. Game theory teaches that a *strategic* response always *dominates* a *sincere* one.

Let me explain this statement, because it is (or it should be) at the center of discussion of all the procedures for making inferences about tastes and preferences. All rational actors are better off if they analyze their potential actions on the basis of the results these are likely to produce, and select the action that will further their

goals. So, if the representative of a country wants a 5% increase in agricultural subsidies whereas all other representatives want to keep these subsidies at current levels, what should be his response to the question 'how much do you want to increase the subsidies?' It depends on his analysis. If he thinks that he is only one representative out of 25 countries, and consequently his opinion is not going to matter in the upcoming discussions, he may tell the truth. If he thinks that asking for 5% will lead to a compromise solution of 1%, whereas if he asks for 10% or 15% he may end up with 2% or 3% as a compromise, he may answer that 10% or 15% is his real preference. If he thinks that 10% or 15% will be considered outrageous by his colleagues, and that even 5% will be judged excessive and generate negative predispositions among the other players, he may respond with 3%. So, the answer that the observer will receive reflects not necessarily the sincere preferences of the actor, but a *strategic representation* of these preferences. How can this problem be resolved? How can preferences be separated from actors' strategic reactions?

In most research papers, preferences are actually assumed. Just as, in economics, businesses are supposed to maximize profits, so, for example, politicians are presumed to maximize votes (or the probability of their re-election) and bureaucrats are supposed to maximize their budgets. An exception to this general rule can be found in the so-called spatial models, where preferences are sometimes approximated by some other indicator (for example, per capita GDP is a proxy for how much countries are willing to spend on subsidies, the size of the agricultural sector is used as a proxy for the size of agricultural subsidies that countries may want, or the ratio of trade to GDP is taken as a proxy for economic openness; see König and Bräuninger, 2004). Preferences might also be directly measured.

Four different ways of measuring preferences have emerged in the literature. The first infers them from the voting records of MPs. The second asks experts to assess these preferences along different dimensions and averages the responses. The third uses opinion polls, and attributes to parties the average of their supporters' preferences or the average of their supporters' beliefs. The most recent method for inferring preferences is to score different texts that parties have produced (electoral programs, parliamentary debates, amendments) on a series of dimensions. The DOSEI project employed three of these four methods; inferring preferences from actual votes was not applied because voting was not used in the Laeken Convention on the Future of Europe. In fact, the president of the Convention, Valéry Giscard d'Estaing, deliberately chose to aggregate and synthesize the preferences without using voting. However, Benoit et al. (pp. 291–313 in this issue) used both the text of amendments and the text of debates to assess the preferences of the actors

who produced these texts. Thus readers hold in their hands the whole array of methods that can be used to infer preferences.

Let me start with the measurement of preferences from voting records (which is not deployed by the DOSEI group and has very rarely been used outside the context of American politics) because it will provide us with information that I shall use in the rest of this paper. In fact, the point of this paper is that scholars on both sides of the Atlantic can learn from each other and incorporate the findings of one group into the other's research.

One can think of different bills (say, *m* of them) presented to a parliament and the voting records of the MPs (consider *n* MPs) as an *m***n* table containing all the information that reflects the preferences of the MPs on the underlying issues. Then, applying a factor analytic technique will reduce the number of underlying dimensions and will allow the different voting issues to emerge, since different representatives will be voting in (more or less) similar groups by issue. For example, in environmental bills the pro-environment MPs will vote one way, and the ones with different priorities will vote differently. Similar patterns will emerge with respect to crime, agricultural subsidies, and so on. The US research along these lines has generated two different outcomes. The first is the scoring of members of Congress from different interest groups on a scale of how friendly they are towards their cause (such as Americans for Democratic Action), the second is the NOMINATE scores produced by Poole and Rosenthal (1997). These two scores are highly correlated with each other, giving high confidence that the location of different members of the House and the Senate on a left–right or conservative–liberal axis can be identified. However, the correlation of left–right positions with other dimensions of the policy space (scores defined by the National Farmers Union, the National Education Association, the League of Conservation Voters, etc.) is not necessarily high. Actually, how high such correlations are gives us important information about the structure of the underlying policy space. If all the measures are correlated, the underlying space can be approximated by a single dimension; if not, one needs a multidimensional model of policy formation.

This is how legislators' preferences are measured in American politics, and these are the data that were used in an important scholarly debate to assess whether *other* factors besides preferences structure behavior in Congress. Whereas most of the literature on Congress at the beginning of the 1990s (Sinclair, 1989; Rohde, 1991) argued that parties had increased their influence, Krehbiel (1993), in an article with the provocative title 'Where's the Party?', disputed the argument by claiming that 'party' as a variable did not seem to explain much above preferences in what was happening in the Congress. The essence of Krehbiel's argument was that, if the preferences of

Democrats and Republicans are distinct, then a vote where all Democrats vote on one side of an issue and all Republicans vote on the other correlates perfectly with party affiliation, but the evidence cannot be used as an argument in favor of the significance of parties. For parties to be significant, the preferences of their members should be overlapping, and representatives should be voting with their party *against* their presumed preferences. Krehbiel took the policy positions from the scores of different interest groups, and tested whether committee composition could be explained by them or required the additional inclusion of a 'party' variable. He demonstrated very little evidence in favor of parties.

Then came a series of articles by Snyder (and his collaborators, see Heckman and Snyder, 1997, and Snyder and Groseclose, 2000) who had the brilliant idea of dividing the votes in Congress into decisive and lopsided ones. His argument was that lopsided votes (where Congress splits 65–35 or more) cannot express anything but the preferences of the members, whereas decisive votes (with more competitive outcomes) may reflect the preferences of parties as well as those of members. Separation of votes in this way led to the discovery of a significant 'party' variable accounting for the difference between the two groups of votes.

With this background of identification of preferences in roll-call votes, the DOSEI group researchers tried to identify the preferences of all the participants in the Convention, as well as of the governments of member countries and other institutions of the EU (the Commission, the European Parliament).

Benoit et al.'s paper in this issue is an attempt to test a model by inferring preferences from documents. The project is in progress and promises to be extremely influential. Different combinations of the authors involved in this project have tested the computerized model, which prides itself on being applicable in every language, with very good results. Here the authors apply it to all EU languages.

The method is disarmingly simple: texts are identified by the distribution of the words used within them. An unknown text is compared with two extreme texts with known properties, and the unknown one is interpolated. For example, if one text uses the word 'war' only once and another uses it three times, a text using the word twice is located somewhere in between the two extreme texts. Estimates from the frequency of use of each word are computed and synthesized, and the final outcome locates the new texts in a space where the two known texts are the extremes.

The results are impressive: high correlations are found between text location and the expert scorings of the authors of these texts. The researchers found evidence of the centrist position of Blair in the election in which he led his party to victory (Laver et al., 2003). Furthermore, they found a high

correlation between texts and expert positions in different settings, including the ones reported here concerning the texts and amendments in the constitutional Convention.

It is very hard to argue with success, and Benoit et al. have an impressive sequence of successes so far under their belt. However, I would like to raise two points that could help us in further investigations. The first is the difference between texts and amendments in terms of correlation with expert judgments. The second is the use of 'words' as opposed to 'concepts' for computerized calculations.

Benoit et al. raise and repeat the point that their results are crucially dependent on the extreme texts that the researcher selects. Indeed, their method guarantees that every text will be located between the two extreme ones and, if the selection is inappropriate, the results will necessarily reflect this inappropriate choice. This is why they divide the different texts produced by the Convention into simple texts and amendments. The first presumably express the positions of the different actors, whereas the second are attempts to organize alliances that would support the proposal. Now, although parliamentary amendments are explicit in this political goal of organizing latent majorities and being successful, there is a significant exception with the current EU amendment texts. The president of the Convention, Valéry Giscard d'Estaing, had explicitly stated that he would not put any text to the vote, but he and the Presidium would try to synthesize the different points of view. As Tsebelis and Garrett (2005) argue, this decision was an ingenious maneuver to appropriate agenda-setting powers and shape the final outcome of the Convention. Indeed, the synthesis did not necessarily have to reflect a specific compromise among different points of view; it just had to claim to do so, and then among all the possible compromises the one preferred by Giscard was selected. Be that as it may, the amendment texts did have as a goal their ultimate adoption and, consequently, were aimed at generating a high level of support. Although support is not a sufficient condition for adoption, it is difficult to imagine that in the EU Convention support was irrelevant.

If this is the underlying policy-making process, Benoit et al. are right to distinguish between texts and amendments, because the first are more likely to reflect the sincere positions of different actors, or their attempts to appeal to their home audiences, whereas the second are more likely to reflect strategic considerations having to do with the other actors involved in the EU Convention. However, if this is the analysis, here is the paradox: the correlations between expert assessments of policy positions and text locations are higher for amendments than for texts.

This finding flies in the face of the results of the American literature on roll-call voting. Remember that lopsided votes (where winning cannot be

regarded as part of the issue) were considered to be an expression of sincere preferences, whereas competitive votes also express the influence of parties (which were exercising their influence in order to win). Here, expert opinions look more like the image produced by strategic representation of preferences than like sincere ones. Why this discrepancy?

One would have to confirm this finding several times before assuming that it is a 'fact' to be explained, but several directions of investigation come to mind. The first one is that there are two levels of European discourse, one aimed at the domestic audience and the other at the European audience, and that 'experts' focus on the second. Further, the distinction that Benoit et al. make between texts and amendments is in fact the distinction between documents aimed at domestic audiences and those aimed at European audiences, and this is what accounts for the fact that 'expert' opinions correlate more with the second than with the first. This finding would open a wide avenue of research comparing discourse for national or international audiences, and the method that Benoit et al. have used in assessing the policy positions of text could be extremely helpful.

A different approach would be to question the validity of 'expert' judgments. As we will see below when we discuss the Dorussen et al. contribution to this issue, the opinions of experts present significantly more variance than one would expect. As a result, simply summarizing the opinions of experts (as has been the case in the literature – see, for example, Laver and Hunt, 1992; Castles and Mair, 1984) can no longer be considered an unquestionable expression of actors' preferences.

The second major issue I would like to raise with respect to the Benoit et al. methodology is their reliance on words rather than concepts. Here is my objection in a nutshell. Consider two parties, with the following programs: the first is the hippy party with the program 'Make love, not war'; the second is the war party with the program 'Make war, not love'. Any analyst would classify these two programs as opposite on the scale of love and war, and would try to locate all other parties somewhere in between, given that each would find some exceptional conditions under which love or war were to be permitted, or required, or excluded.

And here is the distinction that I want to raise. On the basis of concepts, the two programs are opposite in position, because the first approves of love and disapproves of war, whereas the second does the opposite. On the basis of frequency of use of words, however, the two programs are *identical.* Each uses the words 'love', 'war', 'make' and 'not' with the same frequency, albeit in a different order. In arguments, we do use negations and qualifications of different words and this is how we express our opinions. War may be a choice or a necessity (to recall the US political campaign). It may be desirable or

undesirable, or necessary under certain conditions, or for a certain period of time, or to promote particular goals. All these qualifying words are defining gradations in my position on war, they express my definition of 'legitimate' or 'acceptable' war, and it is not the same if one simply counts the qualifiers instead of applying them to the appropriate concept.

I do not think that Benoit et al. would object to these comments. Obviously, including rules so that qualifiers are associated with the corresponding word would complicate the program and would deprive it of an important selling point: its applicability across languages. On the other hand, it would significantly increase the accuracy of its assessment of policy positions.

And here is the essence of the compromise. In debates we tend to repeat the other person's opinion, adding a qualifier or a negation. But do we use such words often in political programs, or manifestos, or legislation? Is it possible that politics operates through issue ownership (Riker, 1993), so that parties discuss only the issues they want to and leave other issues to their opponents? Some preliminary counts would answer these questions. First, identify different qualifying words such as 'some', 'maybe', 'possibly', 'under certain conditions', as well as negations, and count their frequency in different kinds of political discourse. If this frequency is extremely low, then counting words may be an excellent approximation. If not, we would have to move towards counting concepts instead of words.

König et al.'s paper (pp. 269–90 in this issue) is a big departure from all the previous preference-assessing literature. The fundamental assumption of this literature is that what we are measuring is the underlying phenomenon. Critics retort that these measures are contaminated by strategic action, but nobody has actually tried to measure any strategic component of preference revelation. König et al.'s article is the first to do so.

The authors ask what if 'the actor was able to disguise the [policy] position for *strategic purposes*' (p. 280). When they test the DOSEI data, they find that a large distance from the average position of the other national actors and a small distance from the status quo are associated with a lack of policy positions. They interpret this as follows: 'A large distance to MENP [the mean of other national positions] and a small distance to MESQ [the actor's mean distance to the status quo] would confirm our expectation that actors with strategic views on constitution-building explain the occurrence of missing positions' (p. 281).

I need to make two comments at this point – one with respect to the importance of the step, and the other with respect to its risks. First, the path-breaking step of measuring 'strategic behavior' requires an assumption fundamentally different from that in the existing literature. In the existing

literature, what is measured is actual preferences (plus some error term of course). Here, what is measured is the strategic representation of these preferences. Table 1 describes the difference graphically. Table 1 gives two options to actors (to give sincere or strategic representation to their preferences) and two choices to the researcher (to accept the statements of the actors at face value, or to apply the corrective algorithm that König and his colleagues advocate). The traditional literature locates itself in the upper-left corner, whereas König and his associates advocate that the correct location is the lower-right corner.

Having pointed out the big difference that König et al.'s step makes, let me now indicate its dangers. What these authors present is a correlation and we know from other areas that this is (in general) not necessarily evidence of a causal relationship. Additional evidence will be required for the scientific community to reach this conclusion. Yet it is to the credit of these researchers that they have identified this possibility on the grounds of empirical evidence, which is in absolute agreement with the theoretical literature, and have tried to provide means to measure it. What additional evidence is needed in order to attribute missing values to strategic action?

The necessary evidence would be of the sort used in the American debate on party influence. As we saw, it was difficult to distinguish whether parties were *causing* the policy positions or whether there exists a correlation between party and policy positions. What was needed was an independent measure of what could be considered to be unbiased policy positions, and Snyder has identified such a test. An equivalent measure needs to be invented and designed for the EU case. For example, if (along the lines of the existing different discourses at the national and European levels as discussed above) national experts were able to locate the true policy positions or national documents identifying such positions could be found, then the argument of strategic action underlying missing positions would become more convincing.

Assuming that such a crucial experiment to determine strategic missing values is produced, I investigate how to evaluate it. Would König et al.'s suggestion of applying the corrective algorithm always be the correct one?

Table 1 Difference in assumptions and methodology between König et al. and traditional literature

	Accept at face value	Apply algorithm
Sincere representation	Traditional literature	
Strategic representation		König et al. (2005)

As we have seen in Table 1, whereas the traditional approach to preference measurement is that preference revelation is not related to strategic action, and thus no corrective action is required, König et al. assume (and this is not a questionable assumption) that researchers will prefer to apply the corrective algorithm if it is established that missing values are the result of strategic action. In other words, the conventional approach assumes that the probability of strategic action is $p = 0$, and the corrective algorithm should be applied with probability $q = 0$. König et al., on the other hand, assume that $p = 1$ and consequently $q = 1$.

This leads into a discussion of an interaction between the political actors and researchers. And here an interesting question defines this interaction: although researchers want to know the truth, and therefore apply the corrective algorithm if there is strategic representation of preferences by political actors and do not apply it if there is not, what are the corresponding preferences of political actors with respect to the researchers' choices?

One possible answer is that politicians could not care less about what researchers think – after all, how many people read research journals? If this is the case, then König et al.'s suggestion is correct: if we know that politicians apply strategic thinking in the revelation of their preferences, then we should correct for it and $p = q = 1$. Another possible answer, however, is that the reason politicians apply strategic action in the revelation is that they gain advantages from hiding their preferences (König et al. in fact state that '[t]he findings reveal that actors biased by the status quo strategically hide their more supportive positions'). If these advantages were to disappear if their preferences became revealed, then politicians might not be indifferent to the fact that researchers are now able to identify their sincere preferences. After all, this revelation defeats the purpose of strategic action in the first place. So what would political actors do if they prefer to hide their preferences but researchers' preferences for the truth remain the same?

A cat and mouse game between researchers and political actors would ensue. The researchers could apply the König formula and identify sincere preferences (which political actors do not want). Political actors then would prefer to reduce the frequency, q, with which they mask their preferences from 1 to something less, in which case the researchers would reduce the frequency, p, with which they apply the formula from 1 to something less. However, if the researchers reduce the frequency with which they apply the formula, politicians will be able to mask their preferences again, and they will increase their q, which will lead researchers to increase their p. Is there an equilibrium in this cat and mouse game?

The answer is that there always is an equilibrium, and that the exact values of p and q depend on the different actors' preferences. Table 2 presents

an example where politicians prefer to act strategically by e (the difference in their payoffs) when the researchers do not apply the corrective formula and prefer to act sincerely by e when researchers do apply the corrective formula. Similarly, researchers prefer by t to identify the correct preferences of politicians. Regardless of the values of the parameters a, b, c and d, the equilibrium values of p and q (the probability of presenting strategic preferences and the probability of applying the formula) would be $p^* = q^* = 1/2$ in this game between researchers and politicians.

In sum, whereas the traditional literature assumes that politicians reveal their preferences when asked ($p = q = 0$), the König et al. approach (if corroborated by further research) indicates that the information supplied by politicians should be corrected ($p = q = 1$). The assumption that, if researchers can identify and reveal strategic behavior, this will further affect the politicians' behavior leads to a range of intermediate expectations (depending on the payoffs of the actors, with $p^* = q^* = 1/2$ in my example).[1]

Hug and Schulz (pp. 339–52 in this issue) use surveys to identify the public's preferences. These preferences are not necessarily mirrored in governments' preferences. There are many reasons for this discrepancy. The first is that governments are selected on the basis of domestic policies, and they do not necessarily feel that they have to change their preferences to reflect the public's preferences on every issue. The second reason is what Nurmi (1998) has called the 'referendum paradox', according to which electoral districting may cause serious discrepancies between the median voter in parliament and the median voter in the electorate. Third, public opinion may have changed between the time of the election and the time of the survey. Finally, the question has been raised in the literature whether surveys reveal the public's preferences or simply the answers that they give to specific questions (Zaller and Feldman, 1992). Whatever the reason for the discrepancy, a negative assessment of the proposal by the Intergovernmental Conference

Table 2 Game between politicians and researchers

	Accept at face value (in equilibrium, q^*)	Apply algorithm (in equilibrium, $1-q^*$)
Sincere representation (in equilibrium, p^*)	a / $c + t$	$b + e$ / c
Strategic representation (in equilibrium, $1-p^*$)	$a + e$ / d	b / $d + t$

Note: Calculation of equilibrium frequences: $p^* = [(c + t) - c]/[(c + t) - c + (d + t) - d] = 1/2$; $q^* = [(a + e) - a]/[(a + e) - a + (b + e) - b] = 1/2$

(IGC) – the public's preference of the status quo over the IGC outcome – may lead to what Hug and Schulz call 'involuntary defections'. The authors try to identify country candidates for such defections.

The fundamental problem that the researchers face is how to connect the policy space generated by the elites (on the basis of expert judgments of their preferences) with the (less detailed) policy space generated by public opinion surveys. Their answer is brilliant: an n-dimensional space is generated by $n + 1$ points, so, if the same $n + 1$ points get identified in two different n-dimensional spaces, the two spaces can be made to overlap. The application to the DOSEI project is the following. Hug and Schulz identify three points that are identical in both spaces: the status quo, the Giscard project and the IGC outcome. So, they reduce the multidimensional policy space to two dimensions, and then make the three points in each space coincide. As a result they have produced a perfect matching of these two (reduced in dimension) spaces and are able to identify the positions of elites and masses.

From this overlap one can answer a variety of questions having to do with democratic theory. Which countries have the closer correspondence between elite and mass preferences? Is this correspondence a function of the electoral system, of the number of parties in parliament or government, of the coalitions prevailing in government formation? If one takes successive snapshots of public opinion, does the picture vary (which would be the case if the public just answers questions), or is it relatively stable (which would indicate that the public reveals preferences)? This is a remarkably simple framework, opening an impressive number of directions for further research.

The authors choose to answer one specific question that will be central to the ratification process: which countries are likely to create obstacles to the integration process? In their framework, these countries are easily identifiable: they are the ones where the population prefers the status quo over the proposed Constitution (the United Kingdom, Denmark and Luxembourg). The authors report that surveys are more optimistic about the ratification referendums in these countries, but remark that 'voters know very little about the contents of the constitutional Treaty' (p. 348). These issues bring us back to the question of whether voters reveal their preferences in surveys or simply answer the questions asked. If they reveal preferences, there will be some information flow that will lead to more congruence between preferences on issues included in the Constitution and preferences on the document itself. If not, the outcome in these (and other) countries will be anybody's guess.

Dorussen et al. (pp. 315–37 in this issue) try to assess the validity of expert judgments. This is a very important question, given the central role of experts in the attribution of preferences to actors. Expert judgments are still the basis of preference assessments in the literature. It is only off the record that every

now and then one meets an 'expert' who participated in a study and questions the reliability of the study because she was one of the experts. The authors argue that expert reliability (agreement) is a necessary (but not sufficient) condition for the validity (accuracy) of judgments. They control for the number of modalities of a question and for the probability of giving the same answer by chance, and come to the conclusion that the results are quite reliable, with the exception of Hungary (although, given that this is one of the first studies, we do not have any benchmarks). They also find that there is no significant difference in agreement if the person considered to be the leading expert in a country is included in the expert sample or excluded. This is an indirect question mark over the attribute of 'leading' expert. The paper is one of the first attempts to evaluate inter-expert judgments, which is a direction that the literature will have to take in order to assess the validity of its own basis.

Hix and Crombez (pp. 353–76 in this issue) use three different prevailing methods to infer the positions of the different actors in a policy space (exploratory factor analysis, NOMINATE and Optimal Classification). In contrast to results in the American literature which analyzes actual votes and finds that the underlying policy space is one-dimensional (Poole and Rosenthal, 1997), Hix and Crombez find a very high number of dimensions in EU constitutional politics regardless of the method applied. Not only is the number of dimensions very high, but none of them provides an understanding congruent with what the authors call 'intuitive expectations' – a mapping where different countries differ with respect to how much power they are willing to give to the different EU institutions (the Parliament, the Commission, the Council and the Court), which the authors call the 'horizontal dimension', and the federal versus the local level, which they call the 'vertical dimension'. The 'intuitive' understanding of EU politics generates a classification of countries into advocates of a Union of Nation-States (such as France), intergovernmentalists (such as the United Kingdom), Euro-federalists (such as Belgium) and decentralized federalists (possibly the new member states). In order to bring the prior intuitions closer to the factor analytic results, the authors exclude questions that are not relevant to intuitive understandings, and re-analyze the data on this reduced dimension space. In this case, the results are congruent with intuitions.

The issue that Hix and Crombez essentially raise concerns the underlying number of dimensions of the policy space. Is the fact that researchers ask lots of questions and force experts to provide many answers relevant to the overall picture? Or should we rely on an 'intuitive' understanding provided by the 'experts' themselves when they ignore questions that they do not consider relevant? The authors side clearly with the 'intuitive' approach and they disregard data that do not confirm prior 'intuitions'.

However, the issue of the number of relevant dimensions has to be considered very seriously, because this affects actors' strategies and the number (and composition) of possible coalitions. For example, a group of actors may have reached a stalemate in a two-dimensional space, and it is only the introduction of a third dimension that allows concessions to be made to some of the actors and thus makes an agreement possible. Keeping all the questions asked by the DOSEI group and trying to analyze the more complicated underlying spaces, and dropping some of the questions and focusing on fewer dimensional spaces, are both worthwhile tasks.

In conclusion, each of the papers focuses on an issue that is not very well investigated and raises further questions for research. Preferences are essential in politics, and we cannot study politics without them. The assessment of preferences is the subject matter of all these papers. Expert judgments are still the basis of our appraisal of preferences. However, expert judgments should be evaluated for their validity and, as Dorussen et al. argue, we can at least use their reliability as a proxy. In this respect, the DOSEI expert judgments are quite good (and more accurate for the more salient issues). In addition, expert judgments correlate with document assessments (Benoit et al.) as well as (although to a lesser extent) with public opinion data (Hug and Schulz). Each one of these studies will be the basis of a wealth of further research – the Benoit et al. paper because it provides a methodology for extracting preferences from documents; the Hug and Schulz paper because it provides a method for linking the preferences of masses and elites. Because of the strategic implications of the number of underlying dimensions, the dimensionality of the policy space generated by actors' preferences is an important issue to be addressed (Hix and Crombez). Finally, preference revelation is, or may be, a strategic act (König et al.). I say 'is' because at the theoretical level this is always the case; I say 'may be' because König et al. have provided empirical evidence that this is the case.

Taken together, these papers suggest that all the different methods of preference measurement produce similar (correlated) results but we are not at the point that any one of these methods produces the accuracy (validity) of an indisputable benchmark of sincere preferences.

Notes

1 The equilibrium values $p^* = q^* = 1/2$ are not dependent on a, b, c and d. If e and t take different values (as long as they are positive), the outcome will still be a mixed strategy equilibrium, although the values of p^* and q^* will be different.

References

Castles, Francis G. and Peter Mair (1984) 'Left–Right Political Scales: Some "Expert" Judgments', *European Journal of Political Research* 12: 73–88.

Heckman, James J. and James Snyder (1997) 'Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators', *Rand Journal of Economics* 28: 142–89.

König, Thomas and Thomas Bräuninger (2004) 'Accession and Reform of the European Union: A Game-theoretical Analysis of Eastern Enlargement and the Constitutional Reform', *European Union Politics* 5(4): 419–39.

König, Thomas, Daniel Finke and Stephanie Daimer (2005) 'Ignoring the Non-ignorables? Missingness and Missing Positions', *European Union Politics* 6(3): 269–90.

Krehbiel, K. (1993) 'Where's the Party?' *British Journal of Political Science* 23: 235–66.

Laver, Michael and W. Ben Hunt (1992) *Policy and Party Competition*. New York: Routledge, Chapman & Hall.

Laver, Michael, Kenneth Benoit and John Garry (2003) 'Extracting Policy Positions from Political Texts Using Words as Data', *American Political Science Review* 97(2): 311–31.

Nurmi, Hannu (1998) 'Voting Paradoxes and Referendums', *Social Choice and Welfare* 15: 333–50.

Poole, Keith and Howard Rosenthal (1997) *Congress: A Political-Economic History of Roll Call Voting*. Oxford: Oxford University Press.

Riker, William H. (1993) *Agenda Formation*. Ann Arbor: University of Michigan Press.

Rohde, D.W. (1991) *Parties and Leaders in the Postreform House*. Chicago: University of Chicago Press.

Sinclair, B. (1989) *The Transformation of the US Senate*. Baltimore, MD: Johns Hopkins University Press.

Snyder, James M. and Tim Groseclose (2000) 'Estimating Party Influence In Congressional Roll-Call Voting', *American Journal of Political Science* 44(2): 193–211.

Tsebelis, George and Geoffrey Garrett (2005) 'Agenda Setting in the EU Constitution: From the Giscard Plan to the Final Document', Working paper, UCLA.

Zaller, J. and S. Feldman (1992) 'A Simple Theory of the Survey Response – Answering Questions versus Revealing Preferences', *American Journal of Political Science* 36(3): 579–616.

About the author

George Tsebelis is Professor of Political Science at UCLA, 405 Hilgard Ave, Los Angeles, CA 90095, USA.
Fax: +1 310 825 0778
E-mail: tsebelis@ucla.edu