Problems and Solutions in Displaying Results from Alceste

Cheryl Schonhardt-Bailey
Reader in Political Science
Government Department
London School of Economics and Political Science
Houghton Street
London WC2A 2AE, UK
http://personal.lse.ac.uk/SCHONHAR/

INTRODUCTION

For more than ten years, I have used Alceste to study an array of political speeches, legislative debates and policy making meetings. The key advantage in using Alceste—relative to other automated content analysis software—is its ability to capture *themes* (and the associations between and among words, phrases and concepts which this implies) as opposed to more discrete topics (for further exposition, see (Bailey and Schonhardt-Bailey 2010; Schonhardt-Bailey, Lahlou et al. 2010)). But it is not always straightforward to convey the relevance of these themes, particularly as they relate to one another and to the "tags", or identifiers, assigned to portions of the text under investigation. And occasionally, the software itself introduces its own challenges for understanding and conveying the meaning of themes.

For this paper, I seek to address some of the difficulties that I have encountered in understanding the results from Alceste and how I have addressed these in the form of graphics. A basic understanding of Alceste is assumed (see earlier chapters in this volume, or the works cited in the examples below).

To keep things relatively simple, I will examine five specific instances where graphics have been used to overcome what might be perceived as an inherent *problem* in understanding and conveying the program's results. For each, a brief description is given of the data, the problem encountered, the graphical "solution" and a substantive interpretation of the results.

1. The Problem in Visualizing the Significance of Tagged Indicators

A limited number of identifying features are usually given for each speech or subset of text in a corpus—e.g., name, date, role of speaker, party affiliation, and so on. For each class generated by Alcese, any number of these tags may obtain statistical significance, thereby signifying a close association between the tag and the thematic class. But how close is close? What is significant?

Significance for tags is interpreted as the χ^2 value, with one degree of freedom, where:

Statistical Significance (df = 1)	χ^2 value
N.S.	< 2.71
10 %	< 3.84
5 % (*)	< 6.63
1 % (**)	< 10.80
< 1 % (***)	≥ 10.80

But, after obtaining a list of significant tags for our classes, what then? In this first example, the corpora are transcripts from committee hearings in the U.S. Congress, in which the chairman of the Federal Reserve (Fed) testifies and answers questions pertaining to the Fed's Monetary Policy Report (Bailey and Schonhardt-Bailey 2010).

Figures 1 and 2 are graphs which illustrate the extent to which party affiliation, an individual's role (as committee chair or member) and Fed chairman are associated with each thematic class.

[Figures 1 & 2 – about here]

The legend to the right of each graph depicts two groupings of significance—the first from 2 to 50, and the second from 51 to 300. To distinguish the two groups, we size the boxes into small and large, and within each group, the darker the shading, the higher the χ^2 value. As can be seen from the above table of statistical significance, χ^2 values below 10 are less robust but are nonetheless noteworthy. Very high values (e.g., over 50) are, on the other hand, exceptionally robust. Our interpretation does not rely on or adhere rigidly to the specific intervals of these values (e.g., 200 as exactly ten times the significance of 20), but rather to a more relative standard in levels of categories, and particularly the distinction between the two groups of χ^2 values.

The top panel in each figure lists the thematic classes identified in the House banking hearings from 1976-2008, while the bottom panel lists the same for the Senate. From these figures, we can begin to isolate more clearly where committee members of different partisan orientation tended to focus their concerns, and the areas of focus for each successive Fed chairman.

First, both figures show that Members of Congress (MCs) acquire no statistical significance whatsoever for discourse on core issues in monetary policy—e.g., the battle against inflation, the US real economy, monetary aggregates—nor on issues of the US external balance or the world economy. For these issues (which might be called the "guts" of monetary policy), the Fed chairmen dominate the discussion.

A second and related observation is that the discourse of two MCs is unique enough as to create nearly distinct classes for each (we call these our "limelight" MCs)—namely, Senator Riegle and Representative Sanders. The prominence of both is evident from their large chi square significance values in Figure 1a and 1b. We also note that to capture the limelight, a MC need not be in the chair. Riegle's critiques are significant even before he began to chair the Senate Banking Committee, and Sanders' populist anti-Fed remarks are not contingent upon sitting in the role of committee chairman (which he never has).

Following on from this, we also note that—perhaps contrary to conventional wisdom, and committee practice which allows the chair to speak first—the committee chair does not appear to consistently outweigh the rhetorical significance of the members. This is particularly evident for fiscal policy, where the chair scores no statistical significance in either the House or the Senate hearings.

Finally, challenges to the policies and priorities of the Fed tend to come more from the Democrats than the Republicans—as seen in Class 1 (House), noting that Sanders caucuses with the Democrats; and in the Senate, Class 4, along with Democratic Senator Riegle's Class 9. Meanwhile, Republicans are slightly more inclined to discuss banking

regulation, but as this theme received very little attention in the Monetary Policy hearings before the recent financial crisis, it should not be overstated.

Turning specifically to Figure 2, we can see that Fed chairmen are closely associated with particular themes: monetary aggregates for Burns, Miller and Volcker; the fight against inflation for Volcker; the US real economy, the world economy (including trade and current account) and the willingness to explore non-monetary policy issues for Greenspan; and regulation of financial institutions, along with the real economy for Bernanke. Figure 2 provides a concise summary of the priorities of each Fed chairman, but it also encapsulates the changing context of US monetary policy from the mid-1970s to 2008. It depicts the shift from commentary on monetary quantities to the US real economy, with the transition most obviously between the Volcker and Greenspan years.

2. When a Class Disappears from the Correspondence Graph

In our second example, the corpus is the speeches on national security by President George W. Bush and Democratic presidential candidate John Kerry, in the lead-up to the 2004 presidential election (Schonhardt-Bailey 2005). Alceste identifies seven thematic classes, along with their characteristic words and ECUs. Figure 3 is a map of the correspondence analysis of the classes for the corpus. Unfortunately, when the program identifies a large number of classes and the contribution of one of these is relatively small, it occasionally fails to locate the center point for the class with the fewest representative words, which in this case is Class 7. Hence, I have estimated the position of Class 7 from the correspondence analysis of the representative words, as illustrated in Figure 4. That is, using the characteristic words from Class 7, as located from the word distribution in Figure 4, Class 7 is drawn not as a point but as a shaded region, thereby signalling lack of certainty as to its centerpoint.

[Figures 3 & 4 – about here]

The two tags for "Bush" and "Kerry" are superimposed into the same correspondence space, where distance between the classes and tags reflects the degree of association. To the side of the map, we can see that the first two factors together account for just 52% of the total association, with the first factor accounting for about 28%. The relatively moderate degree of association within a two-dimensional space, along with the large number of classes, suggests that Bush and Kerry's discourse on national security contains multiple cleavages.[1] Nevertheless, two of these cleavages account for over half the variation and therefore should be considered prominent.

Figure 3 illustrates a primary dimension along the horizontal axis, namely the same *Global Order – U.S. Specific* cleavage that we saw in the tree diagram. Thus, Kerry's dual critiques of Bush's failings in Iraq and the administration's failure to provide adequate Homeland Security (Classes 2 and 5), Kerry's call to fellow veterans (Class 2), and Bush's War on Terrorism message (Class 4) all fall into the right-hand quadrants, while Kerry's push for nuclear non-proliferation (Class 6) and Bush's dual themes of spreading democratic institutions and fostering economic growth (Classes 3 and 7) fall

into the left-hand quadrants. This is simply another way of observing that the basic content of the seven classes divided into those relating mostly to the U.S. and those concerned with maintaining global order.

Kerry is identified relatively more closely with the U.S. specific issues while Bush is linked more with those relating to the global order. At first glance, this seems somewhat counterintuitive. If Bush appeared to be relatively more concerned with the global order and relatively less with U.S. specific issues, then why did he prevail among the American voters (and, for that matter, become the demon of electorates elsewhere in the world)? This analysis does not, of course, gauge the responses to the candidates' speeches, but if we turn to the vertical axis, we can at least begin to appreciate Bush's appeal to (some of) the American electorate. The vertical axis can be interpreted as separating themes that were expressed in more emotive terms (in the top quadrants) from those expressed using more practical language (in the bottom quadrants). Classes 1, 5, and 6 all fall in the bottom quadrants. All three of these classes were linked to Kerry and all had the common thread of *money* (billions of dollars wrongly spent in Iraq), *personnel* (more specialists required for specific Homeland Security tasks), and *weaponry* (securing nuclear weapons and equipment). In essence, John Kerry was proposing a new way to *Manage the Military,* one that he argued differed considerably from that of the Bush administration. Bush did not significantly engage in this discourse and thus the lower two quadrants of the spatial diagram belong to Kerry alone. (A simpler way to think of this is that Kerry trumped the *Managing the Military* verbal clash between the candidates.)

Meanwhile, Bush was fighting another battle—the battle for the *Hearts and Minds* of the American electorate. In the upper quadrants we find that the common thread between Classes 2, 4, 3, and 7 is an *emotive appeal to shared values.* In Class 2, Kerry invoked his Vietnam wartime experience to capture the hearts of American veterans and their families—a clear appeal to American patriotism. In contrast, Bush tugged at American heartstrings with a different emotive appeal, that of American exceptionalism. By linking U.S. national security with the spread of democratic values and economic growth, and by pitting good against evil in the War on Terrorism, Bush's rhetoric evoked powerful and appealing images in the minds of the electorate. While hindsight is a wonderful thing, the simple conclusion of this full-text analysis is that Kerry's message struck a fairly shallow chord. It appears that Kerry fought and won the battle of logic, but Bush fought and won the battle of emotions.

3. The Challenge in Bridging Words and Deeds

A third case is perhaps less of a "problem" and more of an appeal to researchers to, wherever possible, use textual analysis to marry words with deeds. In this example, the corpus is the U.S. Senate debates of the 2003 Partial-Birth Abortion Ban Act, and a key tag in our analysis is the final vote of each senator on the bill (Schonhardt-Bailey 2008).

Figure 5 presents a map of the correspondence analysis of the classes and tags for the Senate debates on the bill.

[Figure 5 – about here]

Beneath the correspondence map are the percentage associations for each factor, with the first accounting for 44.4% and the second accounting for an additional 32.9%. Hence, a two-dimensional correspondence space accounts for 77.3% of the total variation in the corpus.[2] As an interesting side-note, compared with the much lower percentages from the first two dimensions in Bush and Kerry's speeches and—as we will see in our final example, the seminal speeches of Margaret Thatcher and Ronald Reagan—the percentage obtained for these Senate debates on the abortion bill is markedly higher, i.e., about 25% higher. One might propose (as a testable proposition for investigation elsewhere) that in a debate on a single issue such as abortion, the discourse is more focused than in speeches covering an array of topics (for Thatcher and Reagan) or in speeches on a single topic, national security, but given to different audiences and over a longer period of time (for Bush and Kerry). But, for now, we will limit our discussion to the dimensionality of these Senate debates.

From Figure 5 we can observe first on the horizontal axis a cleavage between the *Personal Experiences* class (where senators relay their own and others' experiences relating to abortion and deformed fetuses) and the other three classes—that is, the *Personal Experiences* class falls at the far left while the other three fall to the right of the mid-point. The "Republican" tag falls in the same quadrant as the *Personal Experiences* class, which suggests that a good number of Republican senators relied on this emotive class in their debates on the bill. That is, Republican senators tended to highlight the gruesome details of the abortion procedure by providing individual stories for an emotive punch.

Second, the "nay" and "yea" tags at the top and bottom of the graph appear to reflect pole positions of the bill's opponents and proponents. Very near the nay tag is the Democrat tag, and both are in close proximity to the *Constitutionality* class. This suggests that opponents of the bill (of which most were Democrats) tended to focus on the issue of constitutionality, arguing that without a health exception the bill would violate the precedent of *Roe v. Wade*, while proponents framed their arguments around more emotive rhetoric (particularly the gruesome nature of the method as detailed in the *personal experiences* theme).

These observations suggest two broad lines of conflict in the debates over the bill. The first and primary (horizontal) dimension can be interpreted as an attempt by the bill's supporters to frame the procedure as uniquely different from other abortion procedures (that is, morally unacceptable), and thereby leave moderate pro-choice senators with no middle ground upon which to stand. The second line of conflict is situated on the vertical axis, and it is this cleavage that appears to underpin the ultimate dimensionality of the roll call vote (from Figure 5), as it pits the opponents of the bill (mostly Democrats) against its proponents (mostly Republicans). The content of this dimension appears to be the controversy surrounding the constitutionality of the omission of the health exception, with the *Constitutionality* theme extending to the very top of the spatial map, while all the three remaining classes are situated in the bottom quadrants. This, together with the close

proximity of the Democrat and nay tags, suggests that while both opponents and proponents of the bill spoke to the theme of constitutionality, Democratic opponents of the bill dominated this discussion.

4. How Do We Transform *Tri-Croisé* Results Into Graphs?

Our fourth case explores the *Tri-Croisé* or Cross-Data analysis in Alceste. Here, we return to the data from our first example—i.e., the testimony of the Fed chairman before the House and Senate Banking Committees (1976-2008) (Bailey and Schonhardt-Bailey 2010). This analysis crosses a tag (name of speaker, etc.) or a single word with the entire text and identifies the strongest statistical associations between the specified tag or word, and other words and phrases in the text.[3] Here we cross each of the Fed chairmen and each of the two major party affiliations with the entire set of House hearings, and again, with the entire set of Senate hearings. This allows us the ability to identify those words and phrases which are most closely associated with Burns, Miller, Volcker, Greenspan and Bernanke, as well as those most closely associated with Republicans and Democrats in the House and Senate hearings. (As a conceptual short-hand, this analysis is akin to holding constant each of our relevant tags.)

For each relevant tag (Fed chairman, party label), the program generates two classes (each with characteristic words and phrases, ordered by chi square significance). One class is unique to the vocabulary of the Fed chairman or party label, and the other class consists of words and phrases that are *least* associated with the tag. We focus here only on the first class, and from that, we examine the top ECUs (which, automatically generated, number 19). Notably, the challenge here is to reduce many hundreds of pages of Alceste print-out into just two graphs.

From a close reading of each set of 19 ECUs, we tally and group them into common categories. While there is of course some overlap between these categories and our thematic classes reported in Figures 1 and 2, our cut into the data here is different, as we are not attempting to use the program to identify themes across the whole corpus, but rather simply to identify vocabulary that is statistically associated with a particular Fed chairman or group of political party members. Moreover, for simplicity, we do not weight or list the ECUs in terms of their χ^2 ranking in the *Tri-Croisé* reporting, but rather treat all the reported ECUs equally. We are therefore applying a structured approach to capturing the meaning of the text.

[Figures 6 & 7 – about here]

From Figures 6 and 7, we can make a number of observations: (1) the distribution in topics between Burns and Miller (in both the House and Senate) is quite similar, which is perhaps not surprising, inasmuch as both were chairmen during the period of high inflation in the 1970s; (2) indicative of his revolutionary shift in monetary policy, Volcker is almost exclusively associated with discourse on money and inflation; (3) both Greenspan and Bernanke devote considerably more attention to the US real economy

(including labour markets) in the House hearings than the Senate hearings; (4) as one would expect with the financial crisis, Bernanke devotes considerable attention to discussing the regulation of financial institutions—but surprisingly, far more so in the House than in the Senate; and (5) the political party divide is far more distinct in the Senate than in the House, with Republicans focusing predominantly on fiscal policy and Democrats on the US real economy (particularly labour markets). This divide still appears in the House hearings, but overall, the discourse is spread across more areas.

5. To 3-D or Not to 3-D?

For our final problem, there is no easy solution. Namely, how do we display three dimensional graphics for publication in journals and monographs in which—generally speaking—the format is black and white, two-dimensional graphics? In this example, our corpus contains the seminal speeches of Prime Minister Margaret Thatcher and President Ronald Reagan, each while in office (Schonhardt-Bailey, Lahlou et al. 2010).

Figure 8 presents a map of the correspondence analysis of the classes (in bold) and tags for the combined seminal speeches of Thatcher and Reagan. The first two factors together capture only about 52% of the variance which is explained (measured by total inertia, which in this case is 80%) in the original correspondence table. [4] Hence, our initial observation is that a two-dimensional correspondence space appears to miss almost half of the sources of variation.

From Figure 8 we can observe first that the horizontal axis mirrors the cleavage between the two groupings of classes in the tree map (*Domestic Economic Policy* and *Protecting Values and Security)*—that is, the *Domestic Economic Policy* group of classes falls in the right-hand quadrant while the other three fall in the left-hand quadrant. As a continuum, the horizontal dimension might range from more political themes on the left and more economic themes on the right. Reagan's tag falls on the side of politics while Thatcher's tag aligns more with economics. In short, the primary horizontal dimension appears to be one of political values and security versus economic policy.

The vertical axis, or second dimension, appears to distinguish the unique language of international politics (*Soviet Threat*) from that of domestic politics and policy. The starkest juxtaposition is with Reagan's *Civil Religion* class, which bridges religious / inspirational values and American society. As *Civil Religion* and *Soviet Threat* fall at the extremes of the vertical axis, this dimension appears to differentiate most clearly the domestic and international aspects of themes relating to politics rather than to economics.

A further feature of Figure 8 is that while the name tags of each leader are obviously constrained to a single spatial location, the venues of the speeches allow insight into occasions during which each leader placed a greater or lesser focus on particular themes. For instance, Reagan appears to have highlighted *Civil Religion* in his inaugural and farewell addresses, as well as in his speeches to conservative (CPAC) audiences. Thatcher tended to focus more on domestic economic policy (i.e., much of it in the form of legislative proposals) in her House of Commons speeches, and more on foreign policy

(*Soviet Threat*) during her speeches at the Lord Mayor's Banquet at the Mansion House in London. As this Mansion House speech is traditionally a forum for prime ministers to discuss British foreign policy, the statistical significance of the Lord Mayor's tag for this international theme might have been anticipated.

While Figure 8 provides some insight into the spatial relationships between the classes and tags, and among the classes themselves, a two-dimensional representation of these data invariably misses nearly half of the variance which is explained in the original correspondence table, as noted above. Moreover, it may seem puzzling that Thatcher's tag is situated in close proximity to the *Big Government* tag for which Reagan's tag is statistically significant. While this is no doubt the result of the overlap between the two leaders on this theme, its spatial location nonetheless seems misplaced. These two features from Figure 8 lead us to investigate a third dimension.

[Figures 8 & 9 – about here]

We know that a third factor explains a further 20% of the variance; hence, an analysis of the data in a three dimensional space offers us a way to capture 72% of the variance explained in the original correspondence table. A color video of this 3-dimensional image may be viewed on-line (http://personal.lse.ac.uk/schonhar/); however, for presentation here, we simplify this image in four rotating "snapshots" (Figures 9a, b, c and d) in which the cell entries represent the clusters of the *top* most representative words for each class. (The on-line version presents the distribution of *all* the characteristic words, illustrated in "balloons".) Figure 9a presents the 3-dimensional space from the same angle or projection as the 2-dimensional correspondence graph—that is, the first and second dimensions are visible and the third dimension is essentially flattened on that plane. The second dimension is fixed in a vertical position—akin to a pole—around which the first and third dimensions rotate in "spokes" at 90 degree angles to obtain the remaining images (as illustrated in the small maps in the top left of Figures a through d). Hence, in Figures 9b and 9d, the second and third dimensions are situated on one plane and the first dimension is essentially flattened on top of that. Figure 9d provides, in our estimation, the clearest view of the third dimension since the classes that appear to drive it fall to the very right and toward the front of the graph.

We find that one class in particular appears to pull the third dimension furthest outward—i.e., *Public Services Reform*—with *Conservatism* contributing to a somewhat lesser extent. *Soviet Threat* also contributes, but as the distribution of its point cloud lies more at a 45 degree angle, its influence is less direct. On the left side of the third dimension lie two distinct classes—*Civil Religion* and *Big Government*—although their positions gravitate more towards the center.

The substantive interpretation of the third dimension appears to be the unique rhetoric of each leader—i.e., the three classes that fall in the right-hand quadrants are statistically associated with Thatcher and the two classes falling on the left are statistically associated with Reagan. And yet, the picture is not quite as stark as that might suggest, since Reagan's tag is also significant for the *Conservatism* theme. What we *do* glean from the

third dimension, however, is a clearer picture of the particularly unique rhetoric of each leader—i.e., *Public Services Reform* for Thatcher and *Civil Religion* for Reagan, with the remaining classes containing varying amounts of overlap between the two leaders.

We have also managed to gain more insight into the spatial position of the *Big Government* theme. Whereas in the two dimensional graph the center of this theme fell in close proximity to Thatcher's *Domestic Economic Growth* and *Public Services Reform* themes, in the three-dimensional version it is situated near Reagan's other theme, as one would expect on a spectrum which divides the distinct rhetoric of each leader. Whereas the two-dimensional graph obscures this finding, adding one further dimension helps to clarify where the language of the two leaders is more distinct.

Adding a third dimension to our analysis has allowed us a sharper way to visualize the themes on which Thatcher and Reagan overlapped, as well as those where their substantive focus and use of language was very different.

SUMMARY

There are many criticisms that are often made of textual analysis by its sceptics: (1) "we could have known that by simply *reading* the material"; (2) "one can't really understand the meaning of documents without also an in-depth understanding of the context in which it is/was written and the mind-set of the author(s)"; (3) "it is all just a fancy counting of words" *or* "the method is black box, so it's all probably meaningless anyway"; (4) "yes, but politicians rarely mean what they say, so why take their words seriously?"; and so on. The general sentiment behind this scepticism is that we cannot glean meaningful results from applying statistical methods to textual data. I disagree. However, the burden of proof is on the researcher to demonstrate by one means or another, that automated textual analysis can offer important contributions to our understanding of any number of areas of interest to political scientists—including deliberation, the policy making process, political rhetoric, and the link between words and deeds. One way to convey the substance of our results more persuasively is through greater use of graphics, which if used effectively, can simplify the message and make it more compelling. The five examples given above are just a small offering as to the possibilities available to us—but perhaps they will spark the imagination.

Bailey, Andrew and Cheryl Schonhardt-Bailey 2010. Deliberation and Oversight in Monetary Policy, 1976-2008. *Southern Economic Association Annual Meeting*. Atlanta.

Bicquelet, Aude 2009. On Referendums: A Comparison of French and English Parliamentary Debates Using Computer-assisted Textual Analysis. *Government Department*. Wivenhoe, University of Essex. Ph.D.

Greenacre, Michael J. 1993. *Correspondence Analysis in Practice*. London, Academic Press.

Schonhardt-Bailey, Cheryl 2005. Measuring Ideas More Effectively: An Analysis of Bush and Kerry's National Security Speeches. *PS: Political Science and Politics* 38(3): 701-711.

Schonhardt-Bailey, Cheryl 2008. The Congressional Debate on Partial-Birth Abortion: Constitutional Gravitas and Moral Passion. *British Journal of Political Science* 38: 383-410.

Schonhardt-Bailey, Cheryl, Saadi Lahlou, et al. 2010. Quantifying the Rhetoric of Ronald Reagan and Margaret Thatcher. *Working Paper*. London, London School of Economics.

---

[1] In total, five factors are identified in the correspondence analysis. Usually, the dimensionality of the system is one less than the number of classes in the profile (Greenacre 1993: 14).

[2] In total, three factors are identified in the correspondence analysis (with the third factor obtaining an eigenvalue of 0.16 and percent association of 22.7).

[3] For a good example of this technique applied to parliamentary debates see (Bicquelet 2009).

[4] In total, five factors are identified in the correspondence analysis. The third, fourth and fifth factors obtain eigenvalues and percent associations of 0.163 and 20.4, 0.124 and 15.5, and 0.10 and 12.5, respectively.

Figure 1: Distribution of Statistical Significance for Each Thematic Class, by Party and Role of Committee Members

Figure 1a: House – All Years

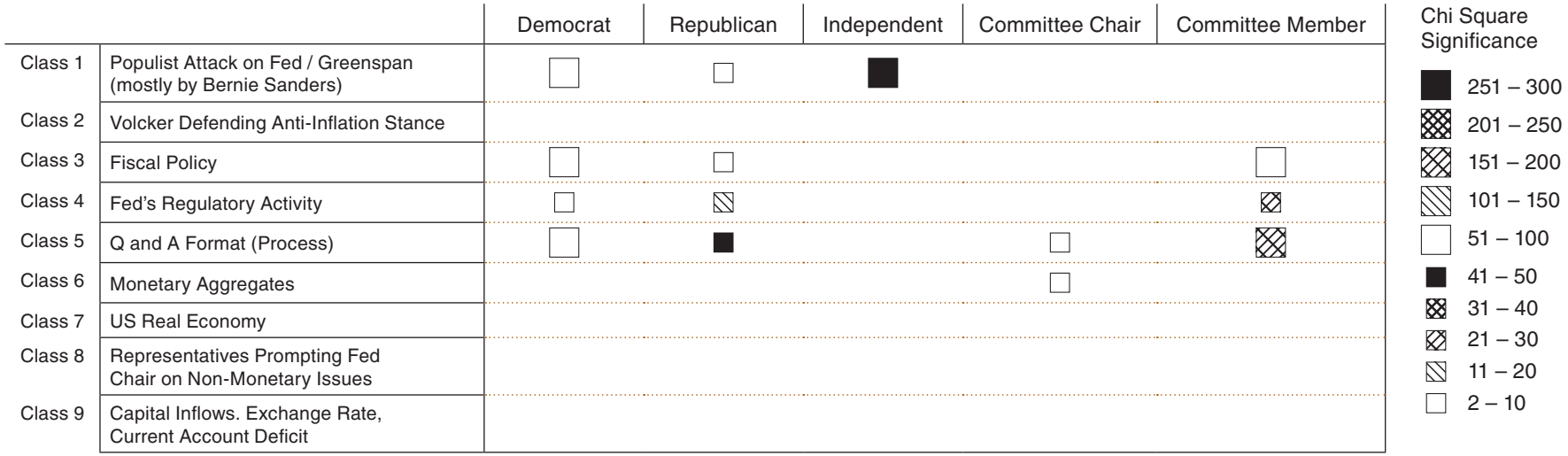


Figure 1b: Senate – All Years

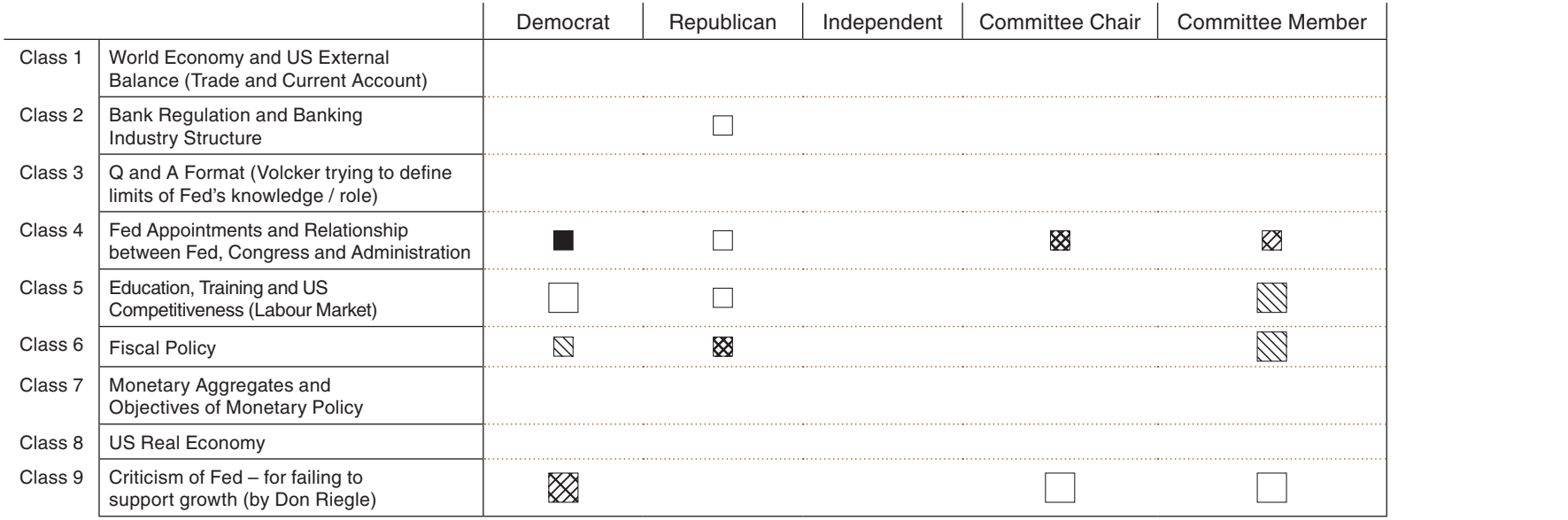


Figure 2: Distribution of Statistical Significance for Each Thematic Class, by Fed Chairman

Figure 2a: House – All Years

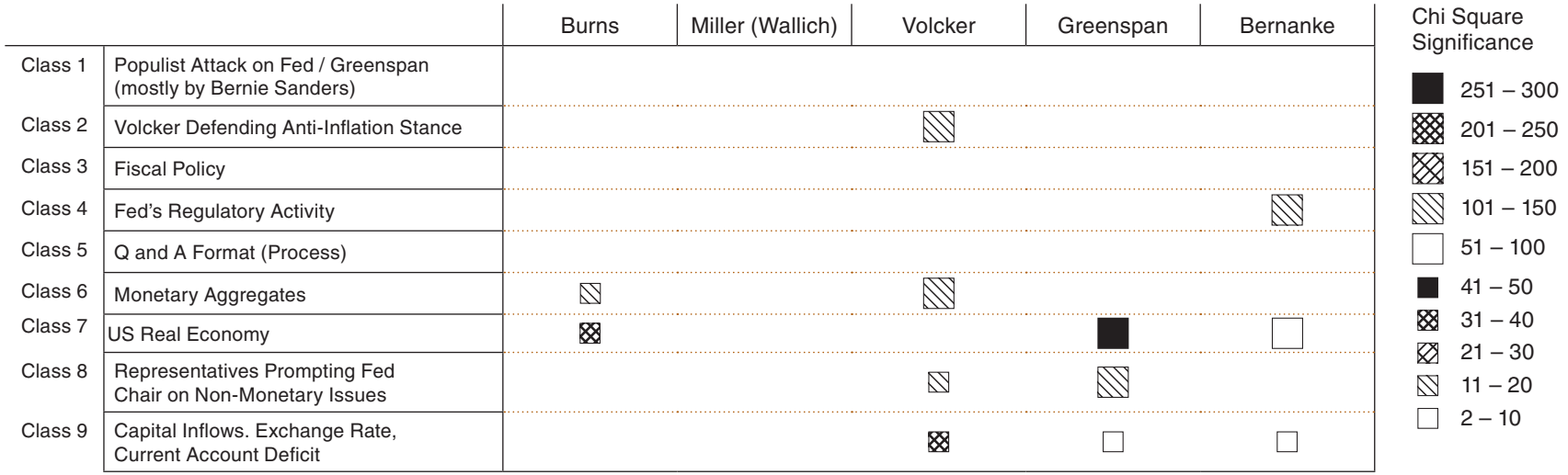


Figure 2b: Senate – All Years

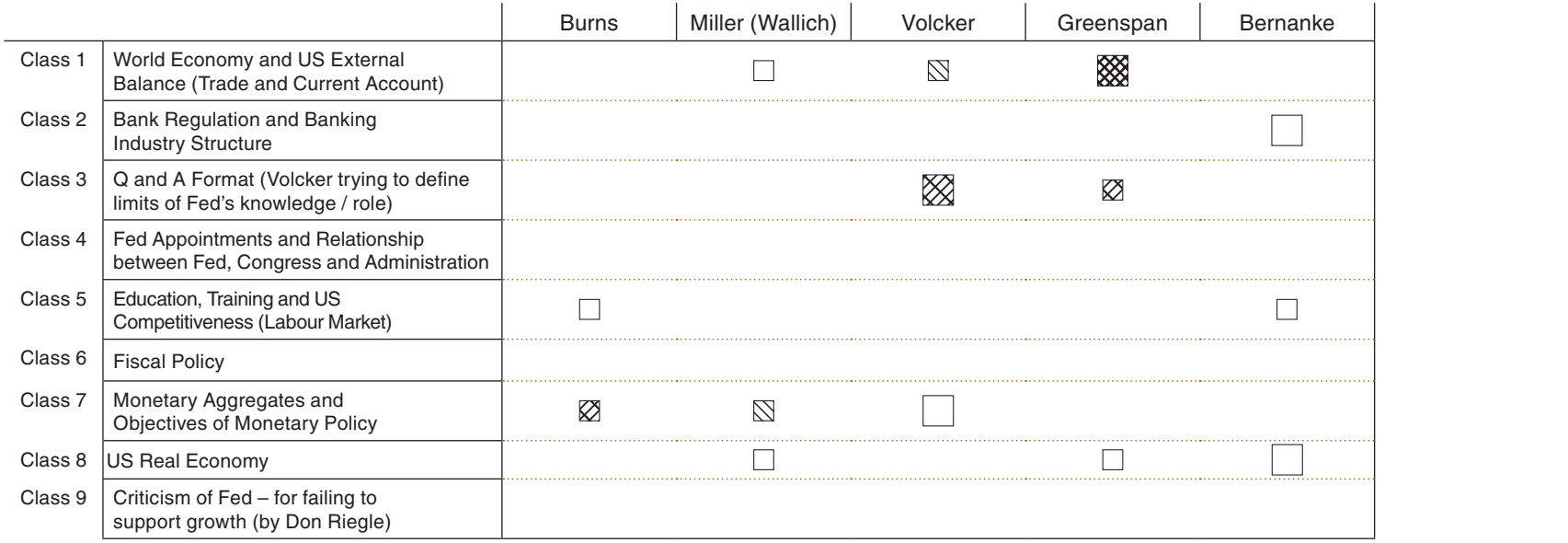
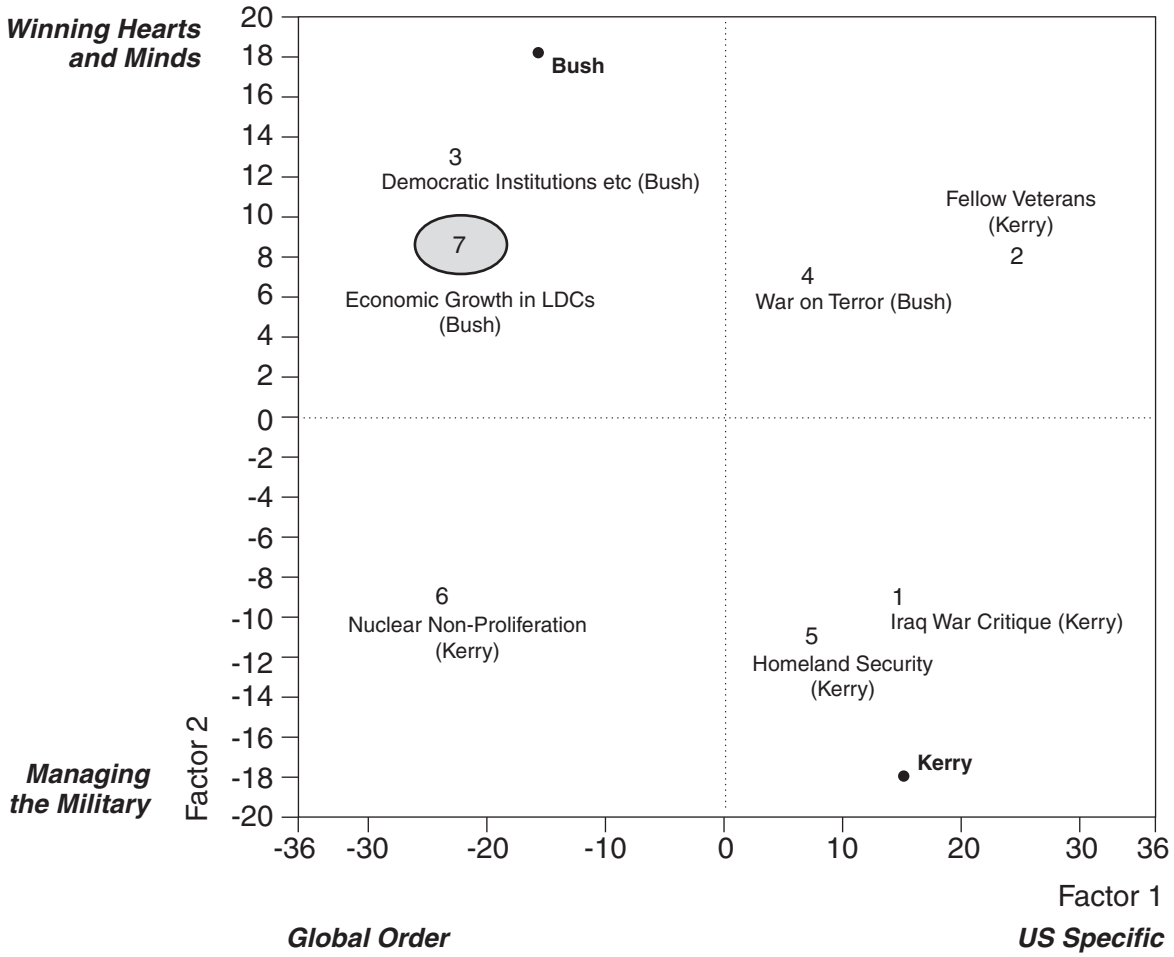
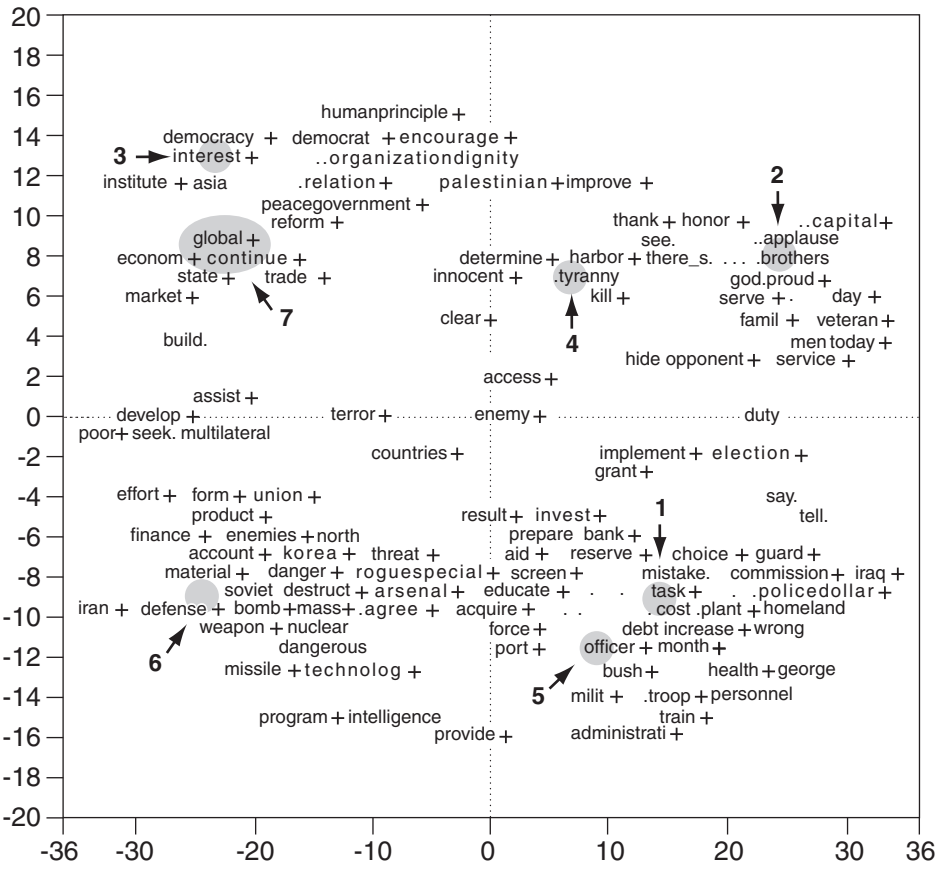


Figure 3: Correspondence Analysis of Classes for Bush and Kerry on National and Homeland Security



	% Association	% Cumulative
Factor 1	28.3	28.3
Factor 2	23.4	51.7

Figure 4 : Correspondence Analysis of Classes for Bush and Kerry on National and Homeland Security, Representative Words

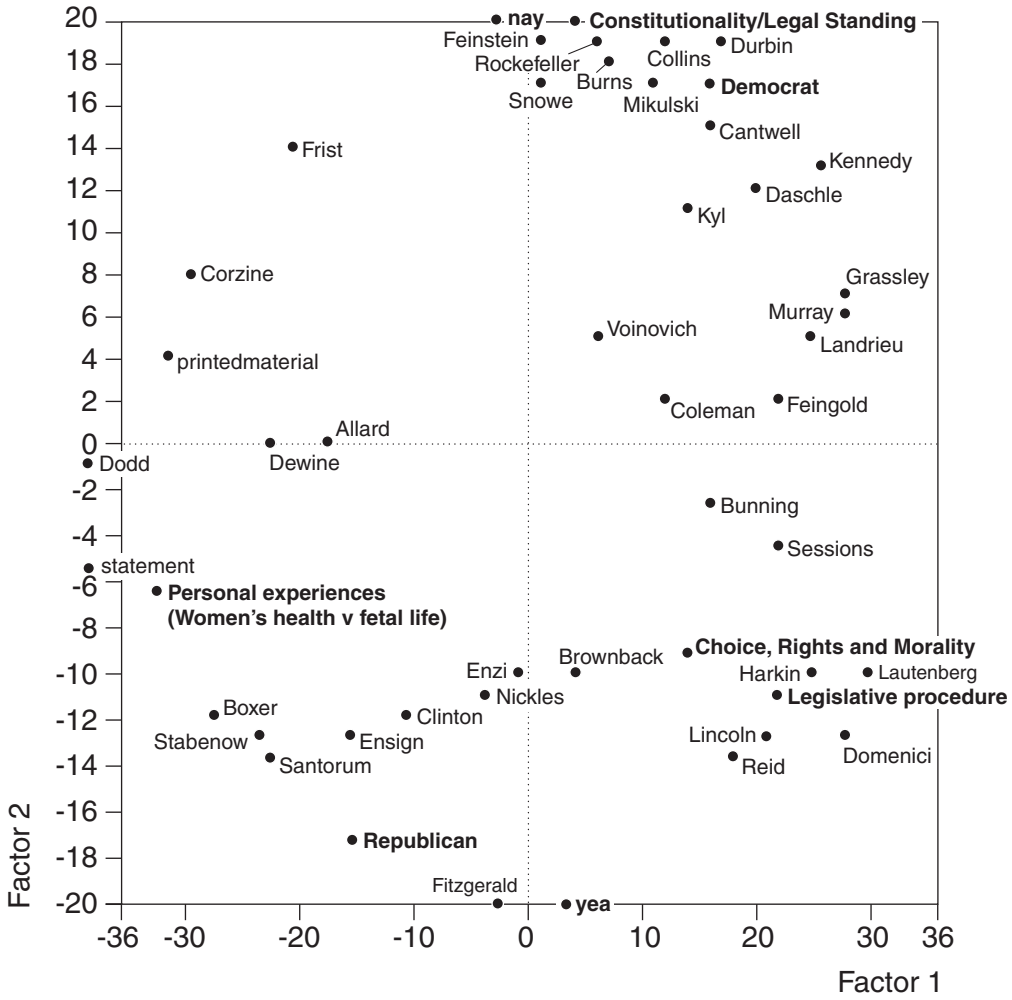


Small dots in the plot signify characteristic words that could not be fitted into the allocated space. These were:

x	y		x	y	
-19	13	particular+	26	8	come.
-19	13	poverty	5	7	september
-19	13	rule+	22	7	appreciate+
-18	13	africa+	25	6	country+
-17	13	region+	8	-9	firefighter+
-18	12	growth	11	-9	miscalculat+
24	10	women	18	-9	bill+
25	10	want+	20	-9	pay.
21	9	pennsylvan+	-11	-10	secure+
22	9	laughter	6	-10	equipment+
20	8	compassion+	8	-10	claim+
22	8	bless.	11	-10	reconstructi
23	8	volunteer+	13	-14	president+
24	8	sacrifice+			

1 **Iraq War Critique** (Kerry)
2 **Fellow Veterans** (Kerry)
3 **Democratic Institutions etc** (Bush)
4 **War on Terror** (Bush)
5 **Homeland Security** (Kerry)
6 **Nuclear Non-Proliferation** (Kerry)
7 **Economic Growth in LDCs** (Bush)

Figure 5: Correspondence Analysis of Classes and Tags from Senate Debates on Partial-Birth Abortion Ban Act



	Eigenvalue	% Association	% Cumulative
Factor 1	0.30	44.4	44.4
Factor 2	0.22	32.9	77.3

Figure 6: Major Themes in House Hearings, by Fed Chairman & Political Party (using Cross Data Analysis)

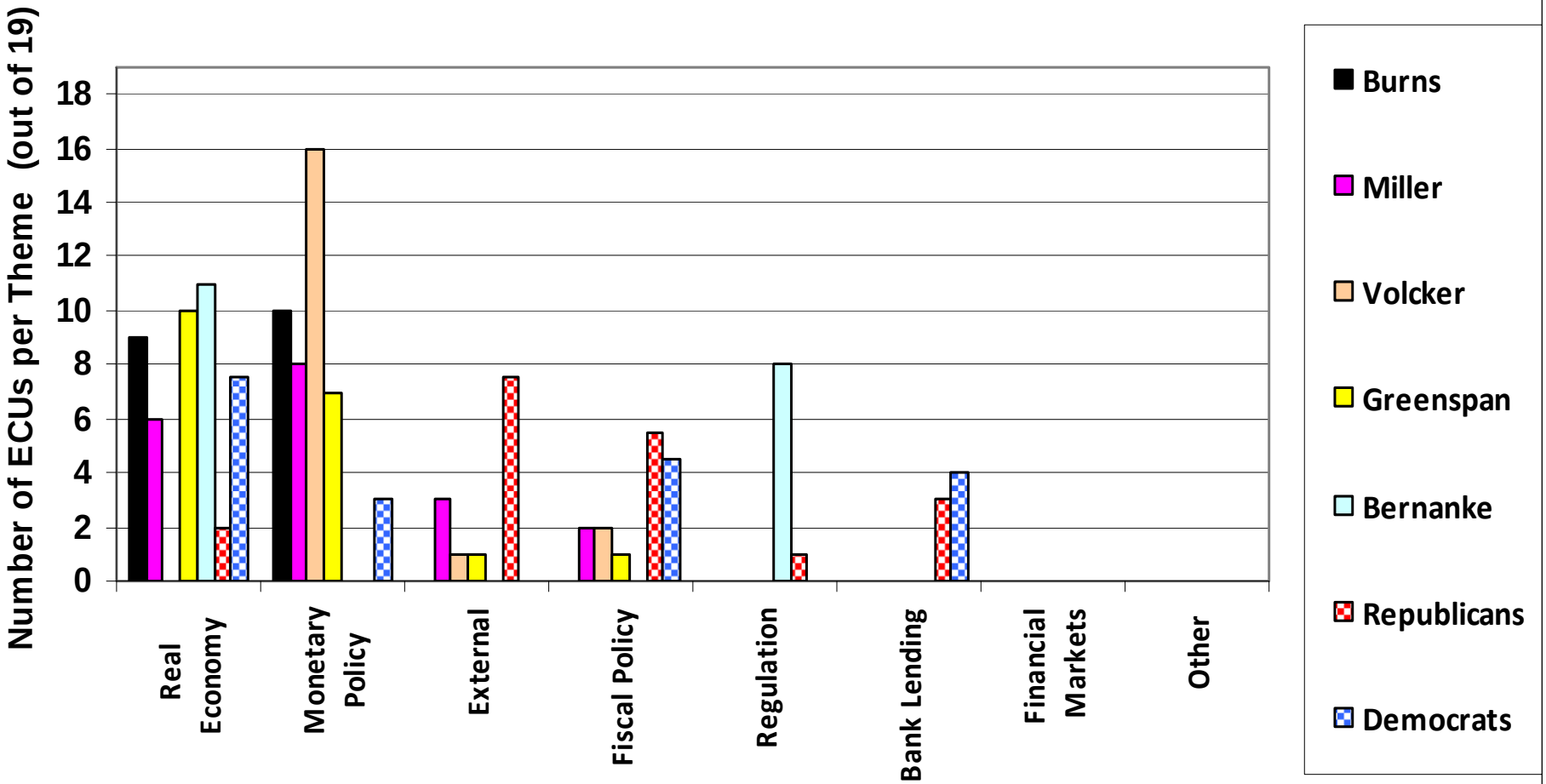


Figure 7: Major Themes in Senate Hearings, By Fed Chairman & Political Party (using Cross Data Analysis)

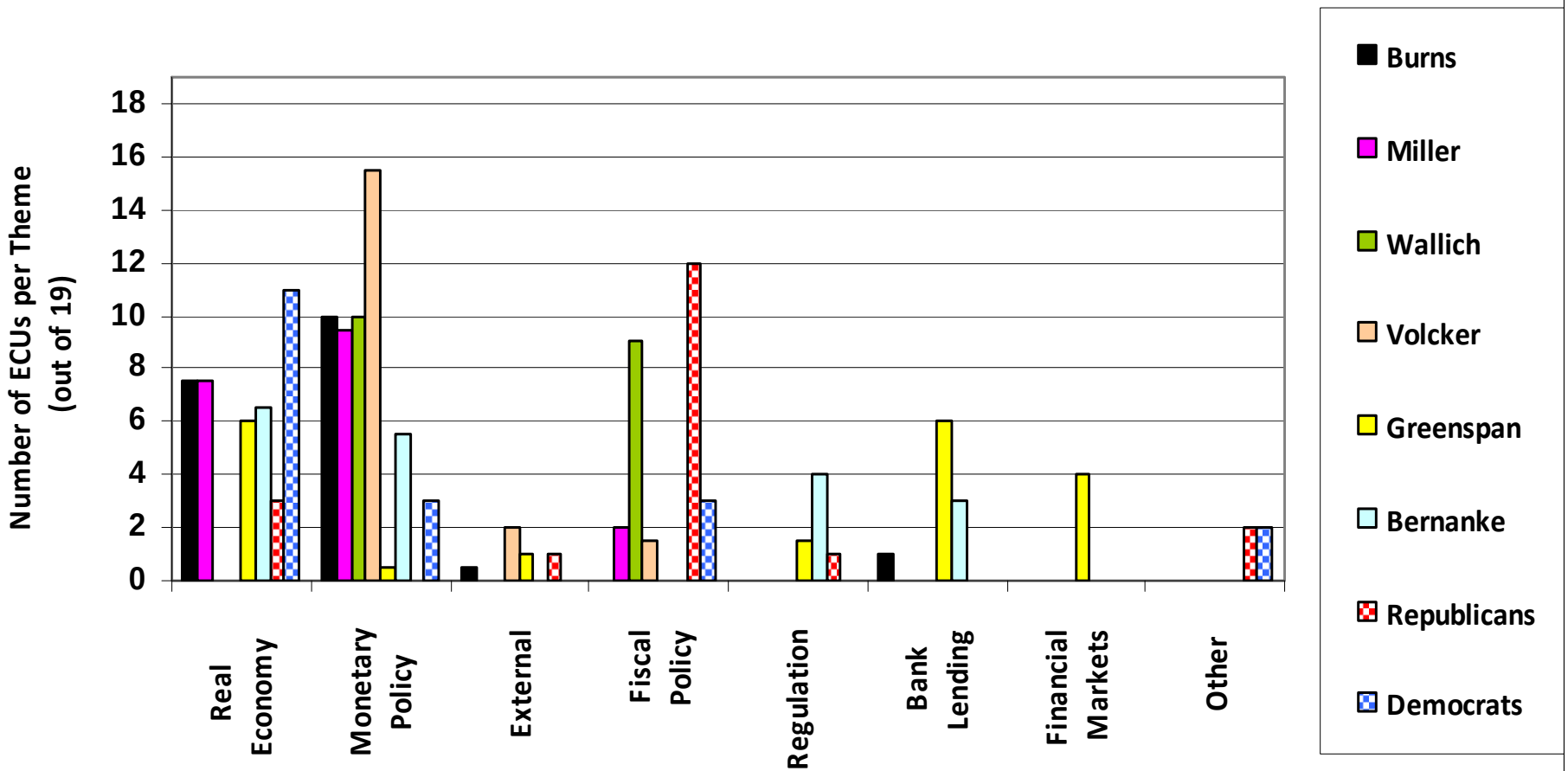
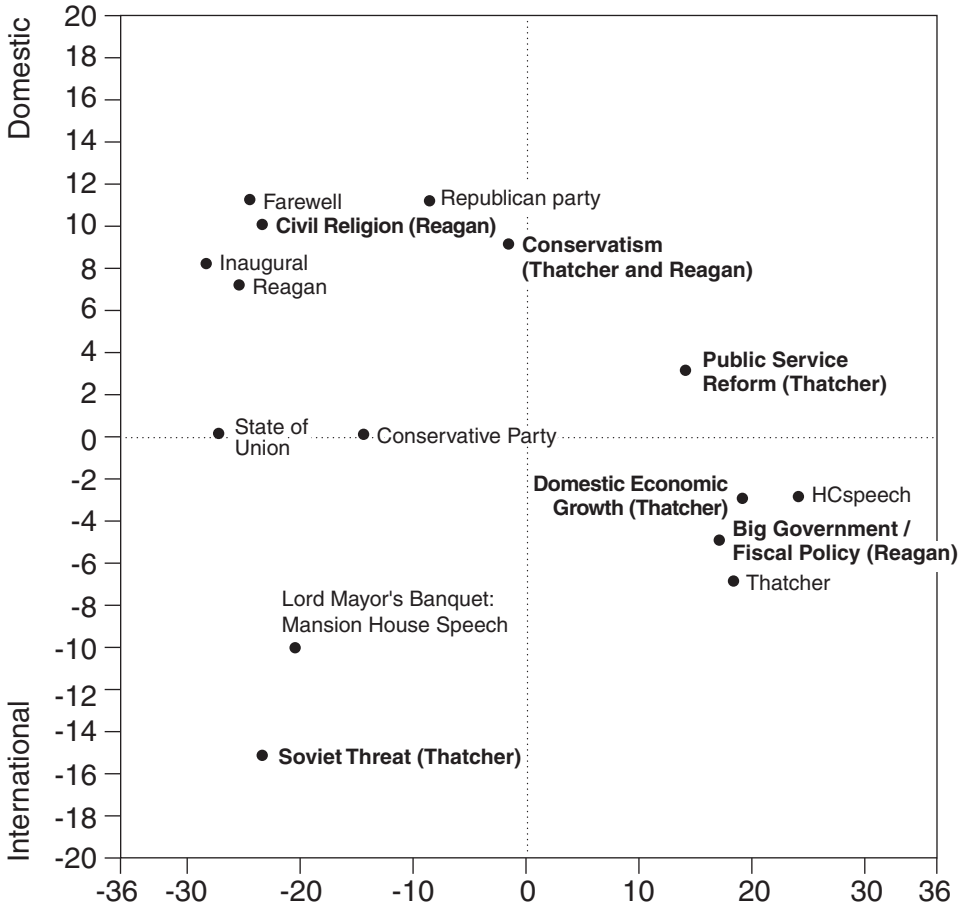


Figure 8: Correspondence Analysis of Classes for Speeches of Reagan and Thatcher

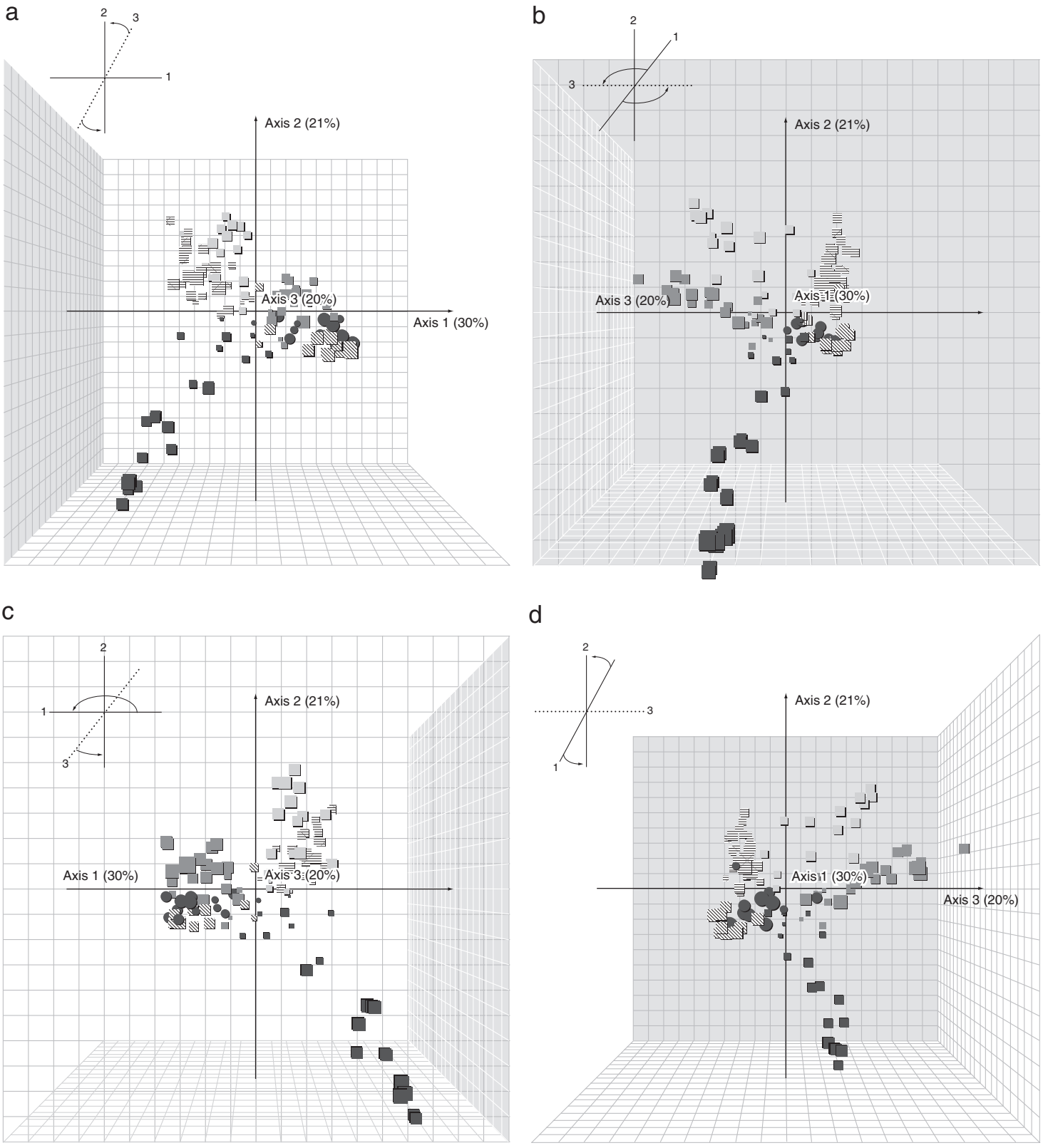


Protecting Values and Security

Domestic Economic Policy

	Eigenvalue	% Association	% Cumulative
Factor 1	0.24	30.14	33.14
Factor 2	0.17	21.47	51.61

Figure 9: Correspondence Analysis of Word Distribution in 3-D*



*The sizes of the cell entries reflects their chi square values.

Soviet Threat (Thatcher)

Public Service Reform (Thatcher)

Conservatism (Reagan & Thatcher)

Domestic Economic Growth (Thatcher)

Civil Religion (Reagan)

Big Government/Fiscal Policy (Reagan)