



**Title:**

Political Involvement and Electoral Competition

**Author:**

[De Sio, Lorenzo](#)

**Publication Date:**

03-16-2006

**Series:**

[CSD Working Papers](#)

**Permalink:**

<https://escholarship.org/uc/item/6xb2g2s9>

**Keywords:**

democracy, democratic, democratization, democratic politics, democratic institutions, political parties, elections, voters, voting, government, politics

**Abstract:**

Recent literature highlights how political knowledge and involvement matter for most political attitudes. Political involvement is here proposed as a second "spatial" dimension, orthogonally complementing left-right ideological positions of citizens. In such a space, a logical quantitative model is developed, hypothesizing that more involved citizens cast votes in a heavily structured manner, strongly connected to their overall ideological orientation, while less involved citizens present higher probabilities of voting for a party that apparently contradicts their self-declared ideological orientation. This logical quantitative model is then applied to data from the Italian National Election Studies 2001 survey. Results confirm the hypothesis, and show that the "competitive area", where voting probabilities are similar for both major blocs, is spatially quite narrow for very involved citizens, and significantly "widens" as the level of involvement decreases. Separate analyses are then carried out for different geo-political areas of the country, with interesting results.

**Copyright Information:**

All rights reserved unless otherwise indicated. Contact the author or original publisher for any necessary permissions. eScholarship is not the copyright owner for deposited works. Learn more at [http://www.escholarship.org/help\\_copyright.html#reuse](http://www.escholarship.org/help_copyright.html#reuse)



# CSD Center for the Study of Democracy

An Organized Research Unit  
University of California, Irvine  
[www.democ.uci.edu](http://www.democ.uci.edu)

Spatial models of voting are probably one of the most known areas of contemporary political science<sup>1</sup>. The “median voter” result by Downs (1957) and Black (1958) has had an enduring influence far beyond the field of axiomatic political theory, also influencing the political debate and practice in many countries, and stimulating survey research.

On the other hand, relevant real-world examples are increasingly putting into question the “median voter” or “convergence” result. The expected result, that is, in two-party systems the two major parties converge in political positions, is often not verified empirically. This also led even axiomatic theory scholars to actually question the applicability of the basic Downs result, by carefully listing which necessary assumptions for Downs’ result are often not met empirically (Grofman 2004). The convergence result then becomes something that is less likely to be observed in empirical results, and new theories are proposed to account for an empirical situation that is more complex than in the basic Downs’ model, thus dealing with multiple issue dimensions, as well as combinations of “proximity” and “directional” approaches (Adams, Merrill and Grofman 2005).

Another approach may though seem as interesting. I propose a two-dimension “political space”, complementing the L-R dimension with the degree of political involvement of individuals. This in turn differentiates such a space from multidimensional conceptions, common both in axiomatic theory (as different policy issues) and in empirical, usually factor analysis-based, results (as latent dimensions of the issue space). Both these latter conceptions mean to unfold the multidimensionality of the political conflict; we are actually focusing on the opposite, that is, the ability of the L-R dimension to catch a large part of the political conflict as perceived by individuals. The second dimension we are looking for must then be something completely apart from attitudes on the politically relevant conflicts or issues.

A reason why combining political involvement with the L-R dimension is *possible* is that, since we assume that the usefulness of the L-R concept stems from its ability to absorb an important part of political conflict, any other dimension (imagined orthogonal) we propose must be as independent as possible from partisan identification and policy attitudes. At first sight, political involvement seems to have such qualities.

---

<sup>1</sup> I am deeply grateful to the Center for the Study of Democracy at the University of California, Irvine, where this article was written in fall 2005. In particular, I gratefully acknowledge the important contribution of Rein Taagepera in the development of the mathematical form of the model applied, and thank Bernard Grofman, Anthony McGann, Carole Uhlaner and Martin Wattenberg for their useful suggestions. The responsibility of how such contributions and suggestions were applied and developed is, though, all upon myself.

This paper is based on data from the ITANES (Italian National Election Studies) 2001 survey. The ITANES research program was conceived and promoted by the Istituto Cattaneo (Bologna, Italy), which is currently responsible for coordinating the program and disseminating data. The ITANES 2001 survey was funded by the Ministry of Education, University and Research. Further information and online data analysis tools are available at <http://www.itanes.it/>.

A reason why combining political involvement with the L-R dimension might really prove *useful* for analyzing voting behaviour (and especially party competition) is that it can be connected with two key variables in electoral analysis.

The first is turnout, which is not only relevant in countries with low turnout rates, but also in countries where such rates are higher, but, since partisan-biased turnout may still have a strong impact on voting results. Neither exploratory analyses nor a specific literature review were yet carried out, but a connection between lack of involvement and voting abstention seems plausible<sup>2</sup>.

The second is the degree of uncertainty in predicting individual party votes. A connection between the degree of political knowledge, interest and several aspects of political attitudes is already made in the theoretical and empirical literature. Less knowledgeable and interested citizens are expected to have less stability of political attitudes; such attitudes are also expected to present a lesser degree of *constraint*, since higher constraint is the result of attitudes deriving by more abstract principles, which in turn can be developed only by more knowledgeable and interested citizens (Zaller 1992; Delli Carpini and Keeter 1996; Van Deth and Elff 2000)<sup>3</sup>. Our basic hypothesis regarding the effect on party competition is that the lesser degree of stability and constraint among less involved voters might increase indifference between parties, and thus make room for a more competitive setting<sup>4</sup>.

The focus of this work is on the second variable we mentioned, that is, the degree of uncertainty in party vote at individual level. Our goal is to run a first test of the hypothesis that political involvement matters for party competition, with less involved voters being more uncertain about which party to vote for<sup>5</sup>. Instead of translating our hypotheses into several functions individually testable with postdictive models, we will develop a predictive, logical quantitative model<sup>6</sup>, which will prove a useful instrument for effectively translating our hypotheses into a testable and potentially falsifiable mathematical relation.

## A Logical Quantitative Model

So far we have introduced the basic features our two-dimensional spatial model should have. First, the “space” we introduced is based on two dimensions with very different features. The

---

<sup>2</sup> About a definition of political involvement, and the relation connecting involvement and political participation, see e.g. Conway (2000).

<sup>3</sup> The theme has also been confronted from a rational choice point of view, mostly concerned with the problem of *heterogeneity* of voters, and of *uncertainty* as the presence of imperfect information. Rivers (1988) proposes a model where voters are heterogeneous in weighing the two dimensions of ideology and party identification; Alvarez (1997) examines how voting decisions and decision processes are affected by “uncertainty”, defined as imperfect information about candidates’ positions. Also, Macdonald, Rabinowitz and Listhaug (1995) test the hypothesis that voters with different levels of sophistication could be described by different spatial models (proximity vs. directional); they find, though, that the evidence does not strongly support such an hypothesis.

<sup>4</sup> Such a result is not trivial, since we could easily have expected the opposite, by claiming that more skilled voters are more likely to change their choice based on rational evaluations, compared to the less involved, deriving their choices by long-standing, possibly inherited, cues. This could especially have been true in historical contexts where long-standing cleavages were still alive.

<sup>5</sup> Our definition of uncertainty only relates to the probability of voting for one of the two parties examined: it is not related to the amount and quality of information that voters have about the candidates.

<sup>6</sup> On logical quantitative models, and on the implications of the distinction between *predictive* and *postdictive* see Taagepera (2005a, 2005b).

first one is what we could call a “ideological” or partisan dimension. It has to deal with the overall ideological orientation of the voter: it is supposed to be able to summarize a large part of the political conflict present in an actual political system, as perceived by citizens. The second dimension is instead divorced from ideology, per se, and should be at best be totally uncorrelated to the first dimension: it should express how much the individual is *involved* in the political conflict summarized by the first dimension. The more voters are interested and knowledgeable in politics, the higher they should score on this dimension, regardless of their ideological position or partisan attachment. In such a space, our first hypothesis is that there is a marked difference among voters with different degrees of involvement. Basically, we propose that, in a two-party or two-bloc system, most involved voters will be, based on their ideological position, sharply differenced in the likelihood of their votes for the two blocs, while least involved voters will be much more uncertain, regardless of how extreme they may be on the ideological dimension. A second hypothesis is that, based on factors such as financial resources, media access and resources, or personal appeal of its leadership, one party may be able to have a systematic advantage over the other, regardless of the spatial position of the voter, thus leading to an overall *bias* of the whole political space.

How to test such hypotheses? Of course there are many possible ways to do that. Our approach is to build a logical quantitative model, attempting to predict the voter’s probability of choosing one bloc instead of the other, based only on her position on the two aforementioned dimensions.

Basic building blocks for logical quantitative models are anchor points, impossible situations, and continuity (Taagepera 2005b). Anchor points happen where our dependent variable will assume values we can be deterministically sure of. An example of such values is the number of different parties that can be represented in an assembly: even knowing nothing about the data, we can be absolutely sure that such number will never go below 1, as it will never go above  $S$ , where  $S$  is the total number of seats available in the assembly. Such remark may seem trivial, but it is actually not, as long as many linear regression analyses can predict incorrect values for these deterministically known cases. Impossible situations are as easy to detect: if our model actually predicts values that are logically impossible, there is no serious reason to keep it unmodified. This can again be seen with a linear regression example. If e.g. a ecological linear regression analysis testing across districts a relation between votes for a party at times  $t_0$  and  $t_1$  yields a negative intercept value, this would actually predict that a party having a low or 0 percentage in a certain district at time  $t_0$  would have a negative percentage of votes at time  $t_1$ . This is an example of an impossible prediction that must be avoided when building a logical quantitative model. The last concept is continuity. This simply deals with the fact that, once anchor points are identified, they must be connected using a continuous function, allowing us to determine values in the middle without assigning any special status to a particular case<sup>7</sup>.

These three criteria combined together actually may define a fairly clear path for finding the appropriate mathematical relation. Once anchor points are identified and continuity is enforced, the choice of possible mathematical forms is usually not really broad.

We will thus move on to identifying our anchor points first. What are values for the dependent variable in extreme cases?

Let’s first remind ourselves that we are dealing with a two-party or two-bloc system. Our dependent variable is the probability that a voter will cast a vote for one of the two parties.

---

<sup>7</sup> From a strictly mathematical point of view, “assigning no special status to any particular case” actually requires a *derivable* function, rather than a simply continuous one.

In our model, we will choose, as reference, to estimate this probability for the right-wing party/bloc. This means that a predicted value of 1 will mean that the voter will certainly cast a vote for the right-wing party, while a predicted value of 0 will mean a certain vote for the left-wing party. We will choose the same reference regarding the ideological dimension: on that variable, we will identify 0 as the leftmost position, 1 as the rightmost. Finally, we will posit that the degree of political involvement may assume a positive or zero value. The zero value refers to people that are completely uninvolved and uninterested in the political debate and conflict happening in that particular country; high values refer to people that have a high degree of involvement.

Once we defined what our variables are, we can identify values for anchor points. Of course identifying such values depends on our hypotheses, and it is actually a way of expressing them. Let's first start by putting involvement aside, and examining what could be the situation for a voter with a "median" degree of involvement<sup>8</sup>. A first anchor point is that, generically speaking, leftmost voters (thus scoring 0 on the ideological dimension) will have zero probability of voting for the right-wing party. Symmetrically, rightmost voters (that score 1 on the ideological dimension) will certainly vote for the right-wing party, then having a probability of 1.

Now let's put continuity in. What will happen between these two extreme points? Our first average guess can be that the ratio between the probabilities of voting for one of the two parties corresponds to the ratio between distances from the two party positions, conveniently identified as 0 and 1. If we then call  $x$  the voter's ideological position and  $y$  the probability of voting for the right-wing party, our guess will translate in the following mathematical relation:

$$\frac{y}{1-y} = \frac{x}{1-x}$$

where the first term is clearly the ratio between the probability of voting for the right-wing party and that of voting for the left-wing party; and the second term is the ratio between the distance from the left-wing party and that from the right-wing party. Voters closer to the right-wing party have a higher probability of voting for it.

It can be easily shown<sup>9</sup> that, in our basic case of average involvement, the formula above boils down to  $y = x$ , thus positing a simple linear relation between the voter's ideological closeness to the right-wing party and her probability of casting a vote for that party. This is basically a standard spatial proximity model.

But it is not by chance that we chose to show the relation in the particular form of Eq. 1. This allows to easily modify the relation to express different degrees of permeability across parties. This is where involvement comes in. We now add involvement, called  $n$ , in the form of a power the second term is raised to. Then

$$\frac{y}{1-y} = \left( \frac{x}{1-x} \right)^n$$

---

<sup>8</sup> Not in a statistical sense.

<sup>9</sup> All proofs are in the Appendix.

The power  $n$  expresses how fast a change in ideological position (ratio) translates into a change in probability (ratio) of voting for the right-wing party. Let's see then how different values of  $n$  actually map to different scenarios<sup>10</sup>.

**Fig. 1- Ideological dimension and right-wing vote probability- different degrees of political involvement**

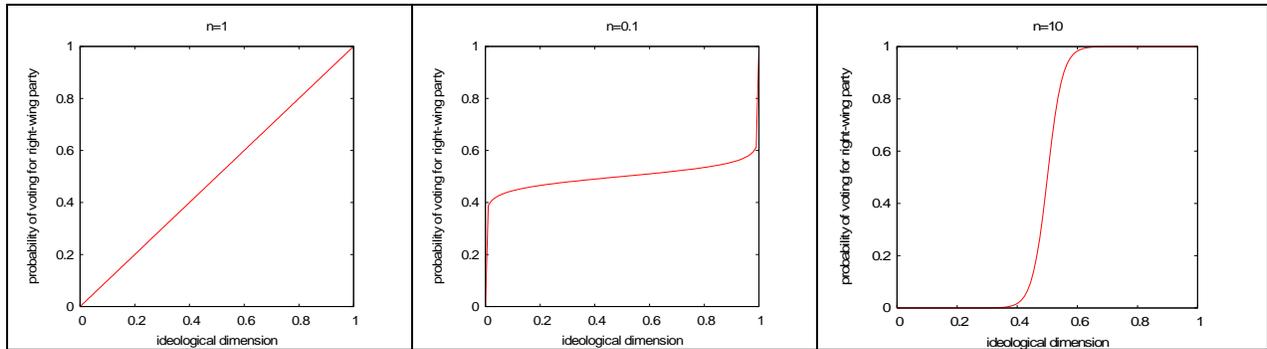


Figure 1 shows how the introduction of  $n$  as an exponent allows us to define different scenarios. The first one, shown in Figure 1a, shows the basic situation for voters of “median” involvement ( $n=1$ ), which was our starting point. In this case, the probability of voting for the right-wing party is simply a linear function of the position on the ideological dimension. This means that two voters divided by a 10% difference in ideological position will always have a 10% difference in their respective probability of voting for the right-wing party, *regardless of where they are ideologically*: the same difference will be present between two almost extreme leftist voters, two moderate leftist voters, two center voters, and so on. This is indeed a scenario where the two blocs are quite “permeable”.

An even more permeable scenario is that depicted in Figure 1b, where we hypothesize voters who are very little involved ( $n=0.1$ ). For such voters, ideological attitudes toward the two blocs make little difference most of the time. A very strong attitude is present only among voters at the far extremes, where probability of voting for the right-wing party still is 0 (for leftmost voters) or 1 (for rightmost). But there is a big difference between a leftmost voter and a slightly less leftist, e.g. one at 0.1. This 10 percentage point change in ideological location is able to immediately increase the probability of a right-wing vote to about 40%, even if this voter is mostly leftist in orientation. And from this point on, the slope of the line is very low: i.e., differences in ideological position make little differences in voting probability for the rightwing party (up to almost the rightmost extreme). These least involved voters act as if they think that “both parties are the same”. As a result, among little involved voters, the two blocs are highly “permeable.”

The opposite situation is hypothesized for very involved voters ( $n=10$ ). In this case, the trend shows that, no matter how moderate they may be, center-left and leftist voters will *never* vote for the right-wing party. This is to hypothesize that, for such voters, their party vote is the result of years and years of reading newspapers, gathering political information, and adjusting

<sup>10</sup> The introduction of  $n$  as a parameter defines a rich and flexible family of possible curves that connect the two anchor points. For  $n=3$ , we obtain what has been known for a long time as the *cube law of votes and seats*, connecting the percentage of national votes with the percentage of seats obtained in a plurality system. Kendall and Stuart (1950) briefly documented the history of the “law” and applied it to the British elections of 1935, 1945 and 1950; the applicability of the “law” was later criticized by Tufte (1973).

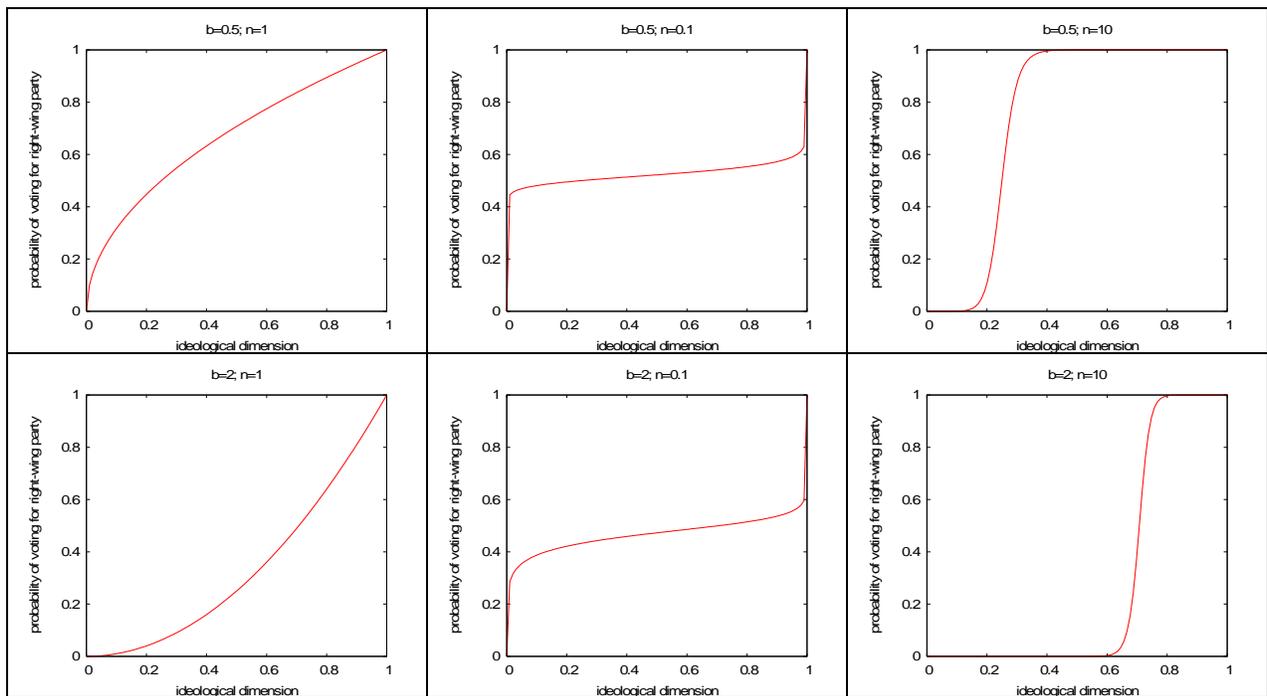
and reinforcing value judgments that are pretty unlikely to be changed. As suggested by the literature on the effects of political knowledge, it is expected that judgments are highly coherent and stable, since they stem from increasingly abstract values. Thus, for highly involved voters, the only permeable region is a narrow center band, where there could really be a 50/50 probability of voting for each party. But as we enter the right-wing zone of the space, involved voters quickly become almost 100% sure of their right-wing vote.

So far, we have looked at the influence of the voter’s degree of political involvement on right-wing vote probability. Let us now introduce the last feature of the model: *bias*. By bias we mean the fact that one of the two parties may benefit from an overall advantage over the other party that derives from e.g. an asymmetry in financial or media resources, or from personal appeal of its leader, etc.

In the case that the right-wing party has this advantage, we expect an overall increase of right-wing vote probability for every voter, *regardless of her ideological position or degree of involvement*. The way to introduce mathematically such a feature is quite simple: we only need to substitute  $x$  (ideological position) with  $x^b$ , raising thus  $x$  to a positive number  $b$  expressing the degree of bias.

This derives from a simple property of power in the 0-1 range. If  $x$  (as in our case) ranges between 0 and 1, no matter the power it is raised to it will be always 0 for  $x=0$  and 1 for  $x=1$ . This is because, no matter the power, 0 raised to any power is always 0, and 1 raised to any power is always 1. This allows us to “play” with that power without altering our anchor points. In this case, a power below 1 has the effect of emphasizing the right-wing range, while a power above 1 has the opposite effect of emphasizing the left-wing range. The effect of such different values of  $b$  can be clearly seen in Figure 2, which repeats the examples of Figure 1 in presence of opposite types of bias.

**Fig. 2 - Ideological dimension and right-wing vote probability – effects of bias at different degrees of political involvement**



The effect of different values of bias is quite clear from the charts presented. To show the mathematical form after accounting for this effect, we simply raise  $x$  to the power  $b$ , and resolve for  $y$ , thus obtaining a full specification of our model:

$$y = \frac{(x^b)^n}{(x^b)^n + (1 - x^b)^n}$$

where  $y$  is the right-wing vote probability for the voter,  $x$  is her ideological position,  $n$  is her degree of political involvement,  $b$  is the overall bias of that election/campaign. Values of  $b$  below 1 actually mean an advantage for the right-wing party, while  $b=1$  means no advantage for anyone, and so on.

### Testing the model

We then proceeded to test the model on empirical data. We used data from the 2001 Italian National Election Studies. The first problem is finding a satisfactory operational definition for the variables involved. Starting from the two independent variables, we generically referred to a “partisan, ideological dimension” and a “apartisan, involvement dimension”. Regarding these two dimensions, many different operationalizations are possible. With respect to the ideological dimension, the standard left-right (or liberal-conservative) self-positioning question is often used, as are factor-analysis derived indices based on policy issues. For the sake of simplicity, and because this is a preliminary exploration, we chose the former.<sup>11</sup>

Regarding the “involvement” dimension, we chose to build a composite index, based on political knowledge and political interest. We summed: the number of correct answers to 5 political knowledge questions<sup>12</sup>; the code of the standard political interest question<sup>13</sup>; the code of a question regarding the frequency of discussing politics<sup>14</sup>. Such sum was then divided by 12 to produce a political involvement index in the 0-1 range<sup>15</sup>.

The second problem is how to operationalize the voter’s probability to vote for the right-wing bloc (where votes for both blocs must sum together to 1)<sup>16</sup>. Since such probabilities

---

<sup>11</sup> Using the standard 10-position L-R self-positioning ruler. One problem with this question is that there usually is a significant share of respondents who are not able to place themselves on the L-R scale, or even deny its meaning. In our sample 568 respondents out of 3209 (17,7%) did not place themselves on such scale.

<sup>12</sup> Respondents were asked to name: the Prime Minister; the President of the *Camera dei Deputati* (House of Representatives); the foreign affairs minister; the President of the Republic; and to state how many members are there in the *Camera dei Deputati*.

<sup>13</sup> Question wording was: “In generale, Lei si interessa di politica molto, abbastanza, poco o per niente?” (generally speaking, are you very, somewhat, little or not at all interested in politics?). Codes range from 0 for no interest at all to 3 for being very interested.

<sup>14</sup> Possible answers were: “Mai; Qualche volta all’anno; Qualche volta al mese; Qualche volta alla settimana; Tutti i giorni” (never; few times a year; few times a month; few times a week; every day). Codes range from 0 for “never” to 4 for “every day”.

<sup>15</sup> Cronbach’s alpha internal consistency test yields a value of 0.606, which is acceptable, given that the index is composed of three items. On the relation among indicators of political interest, frequency of political discussions and political involvement see Van Deth and Elff (2000).

<sup>16</sup> The 2001 elections saw the two opposite *Casa delle Libertà* (House of Freedoms) right-wing bloc and *L’Ulivo* (The Olive Tree) left-wing bloc total respectively 45.5% and 43.7% of nationwide valid plurality votes for the

were not asked of individual respondents, we instead had to compute them at the spatial level, computing percentages of reported votes to the two blocs in every cell of the “gridded” two-dimension space resulting from the crosstabulation of the two spatial independent variables. As a result, we computed such means for each of the 130 cells<sup>17</sup>. It must be also noted that *we only took into account vote to the two main blocs*. This means that the base for computing percentages was the sum of votes to the two main blocs, excluding respondents that voted for other parties or didn’t vote at all. The model thus is limited in that it doesn’t take into account third parties or abstention.

Model testing was performed then on these 130 cases. What we predict is a right-wing vote probability for *all individuals belonging in a particular cell*. (Within the same spatial cell, predicting individual respondents’ vote as other than the mean value in the cell would require adding other variables to the model.)

Once variables were defined, actual model testing requires also defining any *parameters* for the model. The theoretical model’s only adjustable parameter is  $b$ . The degree of overall bias in the campaign is not hypothesized: it must be determined from data, as a result of the estimation procedure, which actually is simply a nonlinear regression of the model we specified.

But, when dealing with empirical data,  $b$  is actually not the only adjustable parameter of the model. Let us see why. We originally hypothesized that the “involvement space” ranges between a value of 0 where voters are practically unable to distinguish between the two parties, and any positive level of involvement, where higher values show that the two blocs’ electorates tend to constitute two perfectly rectangular-shaped, “waterproof” blocs. But our involvement index actually ranges from 0 to 1. First, we need a way to scale such values to a range that could potentially extend to any positive value; but most importantly, we have to account for (and actually measure) the *actual* degree of discernment between blocs for the extreme cases in our sample. Our theoretical model posits that at one extreme there are voters unable to distinguish between blocs, and that at the other extreme there are rigid bloc-voters, where all people left of a thin line vote for one bloc, and all people right of that line vote for the other: but our sample will almost surely present a different situation. Least involved voters will probably still show some difference in their right-wing vote probability based on ideological position; still, highly involved must not necessarily fit into perfectly shaped voting blocks.

Thus we introduced two more adjustment parameters, in order to clearly identify the *range* of the model where our actual sample fits. The first one ( $Z$ ) yields the discernment level of the least involved voters in the *actual* sample. If these voters still show differences on right-wing vote probability based on ideological dimension, such parameter will have a value above zero. The second parameter ( $H$ ) expresses instead the degree in which “involvement matters”. Thus,  $H=0$  would mean that both the least involved and the most involved have the same discernment level  $Z$ . If, on the contrary,  $H$  assumes a very high value, this means that involvement makes a great difference regarding the discernment level between the two blocs. By comparing values of these parameters we are then able to express evaluations about our sample, and also comparison among different samples, as we will see later.

---

*Camera dei Deputati* (upper house). Italy’s 75%-plurality mixed-member electoral system then awarded the *Casa delle Libertà* with a 58% seats majority in the *Camera dei Deputati*.

<sup>17</sup> The involvement index had 13 possible values ranging from 0 to 12; the L-R self positioning variable had 10 possible values ranging from 1 to 10.

Thus we will posit that:

$$n = Z+Hi$$

where  $n$  is the involvement level in our abstract model,  $i$  is the actual involvement level measured by our index on a 0-1 scale,  $Z$  (a parameter to estimate from the data) is the discernment level of the least involved in our sample,  $H$  is the multiplicative parameter that expresses the difference in discernment level between the least and most involved in our sample.

We then substitute  $n$  with  $Z+Hi$  in our model, and obtain an empirically-testable version:

$$y = \frac{(x^b)^{Z+Hi}}{(x^b)^{Z+Hi} + (1-x^b)^{Z+Hi}}$$

where the adjustable parameters to estimate are  $b$  (amount of overall bias favoring one of the two parties<sup>18</sup>),  $Z$  and  $H$  as seen above.

## The Results

The parameters for the model were estimated on the aforementioned ITANES 2001 data with a standard non-linear regression analysis: the results of the estimation for the whole country are presented in Table 1.

**Tab. 1 – Parameter estimate for Italy 2001 (whole country) – ITANES 2001 data**

| Parameter         | Estimate | Std. Error | 95% Confidence Interval |             |
|-------------------|----------|------------|-------------------------|-------------|
|                   |          |            | Lower Bound             | Upper Bound |
| B                 | .912     | .012       | .888                    | .936        |
| Z                 | 1.145    | .345       | .461                    | 1.828       |
| H                 | 9.953    | 1.974      | 6.046                   | 13.859      |
| R-squared = 0.940 |          |            |                         |             |

Thus the estimated model for our sample is as follows:

$$y = \frac{(x^{0.912})^{1.145+9.953i}}{(x^{0.912})^{1.145+9.953i} + (1-x^{0.912})^{1.145+9.953i}}$$

<sup>18</sup> Values of  $b$  are actually a little awkward, since, despite the fact that on both  $x$  and  $y$  higher values have a right-wing meaning, it assumes values below 1 for a bias favoring the right-wing party, and values above 1 for the opposite. Such problem could be resolved by putting  $1-b$  instead of  $b$  in the model, but I didn't want to introduce further complications.

which allows us to directly predict  $y$  (probability of right-wing vote) based on  $x$  (left-right self-positioning) and  $i$  (level of political involvement).

The first interesting result is the goodness of the fit. R-squared has the high value of 0.940, meaning that the fit of the model is fairly good. It must be noted, though, that we are working with a probabilistic model. This has basically two meanings: one operational, the other more broadly theoretical. The operational meaning is that what we are actually fitting are not individual votes or individual probabilities of voting, but “cell probabilities”; actually, counts of votes in cells<sup>19</sup>. The theoretical meaning sets instead the context for our model. We are not trying to predict individual voting behaviour, which the literature has shown being dependent, in different contexts, on a vast array of social, demographic, economic and psychological, individual-level variables, which on the other hand account for a large part of the individual’s ideological self-positioning. We are instead looking at those aggregate, institutional features that shape the political space (ability of parties to integrate the less involved into the main political conflict; degree of knowledge and discernment among the least involved, *bias* induced by disproportionate amount of resources, etc.) and that then should appear as detectable “spatial regularities” that our model wishes to highlight. Beyond this point, individual-level variables are inevitably needed to explain individual votes. These caveats given, the fit of the model seems remarkable, meaning that our hypotheses seem to catch (and to predict) quite well the “spatial regularities” in our sample.

Moving to the estimated parameters, they give us really important information. The first one is  $b$ , that is, the amount of overall “bias” that we observe in the whole distribution, regardless of the spatial position of voters. The value estimated is 0.912, which is a value significantly lower than 1, meaning that the right-wing bloc had an overall higher attractive power<sup>20</sup>, thus having a higher ability of reaching voters on the whole space, regardless of their spatial position, being able to “capture” even some leftist voters. We will soon see graphically the meaning of such bias.

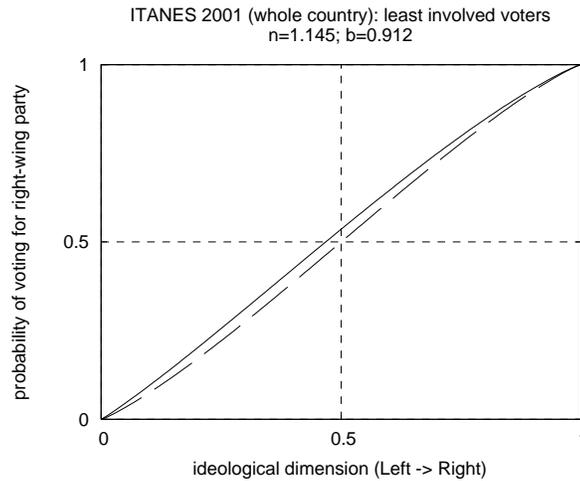
But let us now introduce also the “tuning” parameters regarding the level of political involvement. The value of  $Z$  is 1.145: fairly above 0, and even slightly above 1. This shows that in our national sample, even among the least involved voters, ideological position is far from irrelevant for choosing between the two major blocs. There is an almost linear trend relating to ideological position. Yet a good degree of “permeability” is present: even voters with a clearly defined ideological position far from the centre retain a significant probability of voting for the opposite bloc (Figure 3).

---

<sup>19</sup> However, a binomial logistic regression model, using the predicted probability  $y$  as the only independent variable, correctly predicts 81,1% of the cases.

<sup>20</sup> It is though not a surprising value, since it still ranges in a 0.9-1.1 range, as we might expect.

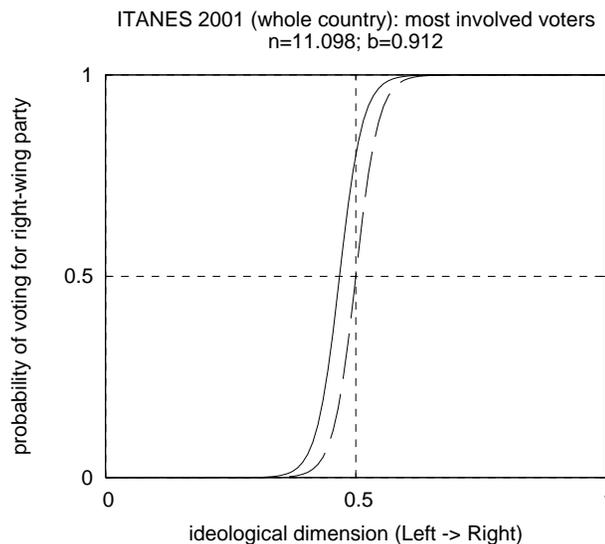
**Fig. 3 – Estimated model: L-R self-positioning and right-wing vote probability among the least involved (dashed line show theoretical absence of bias)**



The figure shows the high degree of permeability between the two blocs among the least involved voters; it also shows a small importance of bias. The two grid lines show the centre point: in absence of bias, the probability curve would be the dashed curve, passing from the centre of the square: in absence of bias, the probability curve would be the dashed curve, thus correctly predicting a 50% probability of voting for any of the two blocs for voters lying exactly in the centre. This behaviour is altered by bias: the solid curve (actually, the best fit derived from the data), shows that the 50% probability line slightly moves to the left. The actual gain in right-wing vote probability deriving from bias is thus the area between the two curves. The effect is however extremely limited.

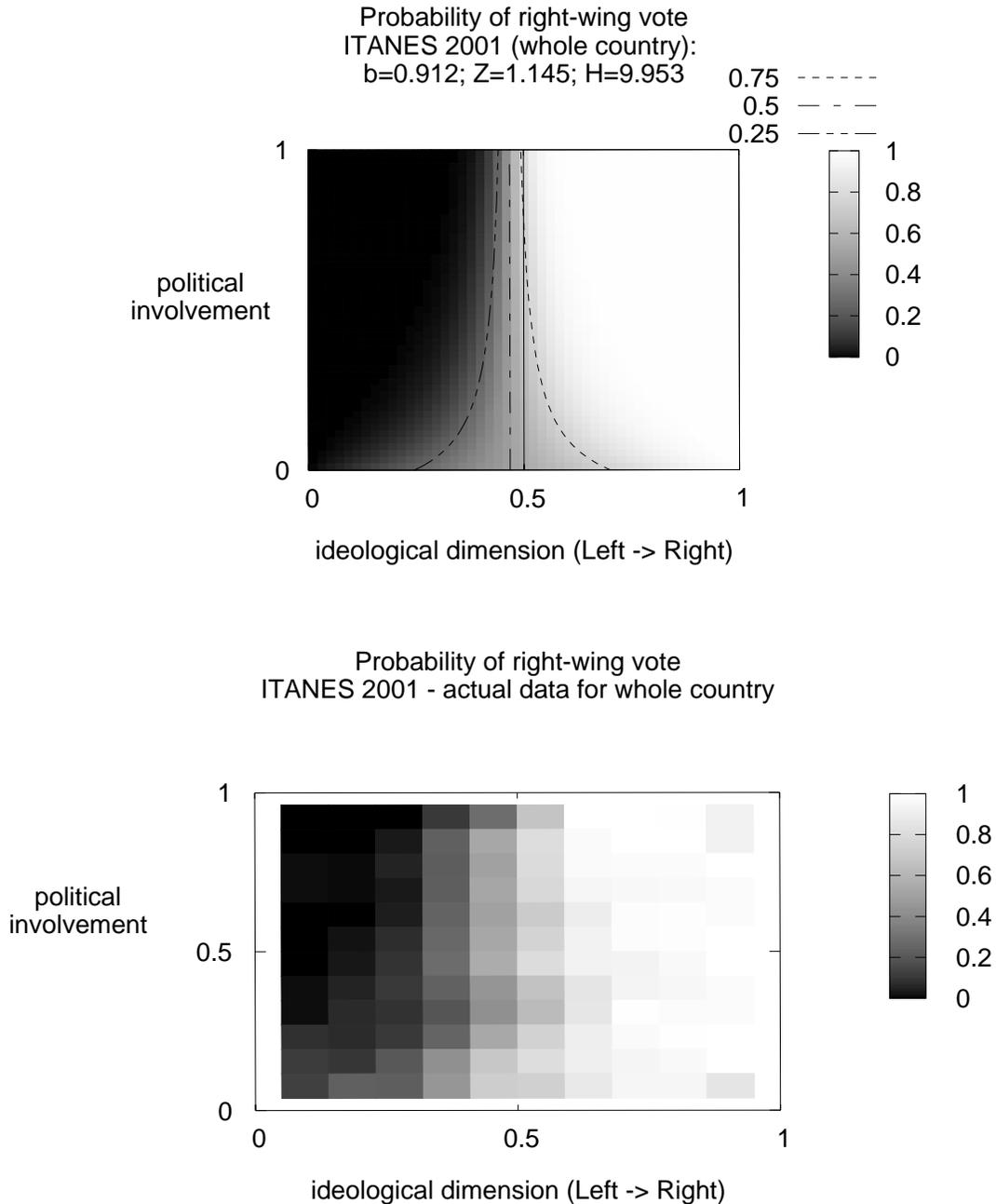
For most involved voters, we have a different situation regarding permeability. The value of  $n$  is much higher, being  $Z+H=11.098$ . This means that the degree of permeability is much lower among these voters. This can be clearly seen in Figure 4.

**Fig. 4 - Estimated model: L-R self-positioning and right-wing vote probability among the most involved (dashed line show theoretical absence of bias)**



The degree of permeability is much lower, since for most part of both camps the predicted probability of voting for the other bloc is 0. Also the influence of bias is still quite limited, since the overall increase in right-wing vote probability coming from bias is again a narrow region between the two curves. The two figures show the trend in the two extreme cases. We summarize the whole trend for all possible degrees of involvement, as predicted by the model, in Figure 5. A similar figure is included, based on actual data.

**Fig. 5 - L-R self-positioning and right-wing vote probability (shades of gray) – Estimated model (above) and actual data plot (below)**



The figure shows different probabilities of right-wing vote in different shades of gray. It also shows the lines of 25%, 50% and 75% probability. The two basic hypotheses of the model, with the actual value of parameters estimated based on data, are easily recognizable. Bias is present, so as to give the right-wing party an overall advantage regardless of spatial position. This is visible by the fact that the 50% probability line is slightly shifted to the left. But also the degree of involvement makes a difference, since the 25%-75% “uncertainty band” is much wider among the least involved than among the most involved. As a first cut, we note three main points:

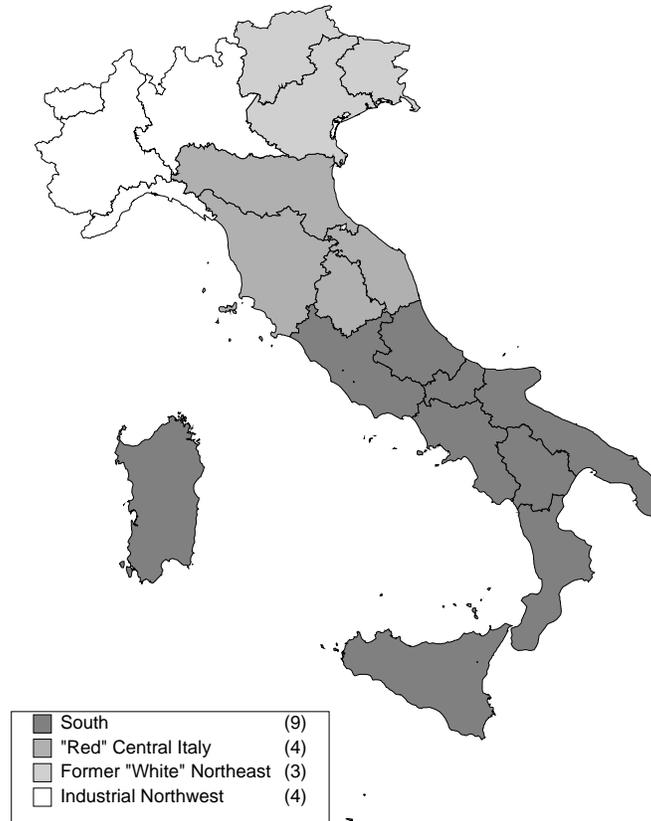
- 1) The fit of the model is fairly good: the hypotheses seem to hold not only qualitatively, but also quantitatively, since the value of R-squared is very high;
- 2) A slight amount of bias seems to be present: this shows that in the 2001 general elections, Italy’s “Casa delle Libertà” right-wing bloc was actually able to draw a certain amount of votes among people who actually positioned themselves as leftist. This effect is however less important than expected. A possible explanation could be random error, or a not yet complete specification of the model: for example, a stronger influence of bias among the least involved could be added. A more radical problem could though be present, that is, the fact that vote “swing” also had the effect of changing individual self-positioning: thus the ideological dimension might better be measured differently, e.g. with issue-based indices.
- 3) There seem to be a significant difference among most-involved and least-involved voters: the latter seem to be less clearly oriented in choosing between the two blocs based on the voter’s ideological position. It must be noted though that an indifference between the two blocs regardless of the respondents’ ideological position is not observed, even among the least involved.

The above considerations were for Italy as a whole. We now proceed to a comparison among different parts of the country. The geographical subdivision we will use corresponds almost completely to that outlined by Galli (1966), distinguishing an “Industrial Zone” (north-western Italy), a “White Zone” (north-east), a “Red Zone” (central Italy), and the remaining South (see Figure 6).

Such subdivision was originally drawn from several self-reinforcing historical factors, whose emergence can in some cases be traced back to differences in social, economic and institutional structures from the Middle Ages on (Cartocci 1987; Putnam 1993). As a result, these four different parts of the country followed different paths also after Italy reached national unity in the second half of the 19<sup>th</sup> century. While the northwestern region rapidly experienced industrial development around the three main industrial cities of Milan, Turin and Genoa, the two still mostly agricultural regions of the Northeast and of Central Italy reacted to the agricultural crisis of the 1880s by developing a dense fabric of cooperatives and mutual relief societies, mainly aimed at exploiting collective action in order to dampen the effects of the agricultural crisis. Such fabric had different common denominators in the two areas. In the Northeast the common identity allowing such development was religion, in the form of a strong devotion to the Catholic Church, fostering association partly also opposed to the anti-clerical liberal State of the post-Unity season; in Central Italy, Socialist values were already present, partly as a 19<sup>th</sup> century development of inherited, middle-age born, minor arts’ workers guilds and associations, still partly alive at the end of the 18<sup>th</sup> century. Both these fabrics eventually became actual “subcultures”, this term relating to the presence of a value and belief system radically alternative to the dominant culture; a system almost hegemonic in its geographical area,

and nested on a associative and cooperative network who was able to socially integrate individuals in most aspects of their everyday life.

**Fig. 6 – Geographical zones**



Fascism was not able to completely eradicate the two subcultures, which flourished even more with the new Republic after the Second World War, as they became a major source of political support respectively for the Christian Democrats and the Communist Party, Italy's two major parties up to the 1990s (Galli 1966). The gradually increasing secularization of society from the 70's on led to a rapidly decreasing strength of Catholic identity, which in the Northeast was coupled with radical changes in social and economic structure. Both factors, along with the collapse of the Christian Democratic party in 1993, led to a virtual end of what was called the "White subculture" in the Northeast. This was not mirrored in a parallel end of the "Red subculture": Central Italy has experienced a continuing strength of the Communist Party and of its successors after 1991, and the associative fabric is in part still alive, though after experiencing profound changes. As a result, the part of central Italy known as the "Red zone" has retained features of its old identity, along with a strong degree of party identification, which during the 80's instead was steadily decreasing in the rest of the country (De Sio 2006). Compared to these developments, the South had and still has a much less clearly defined political identity: the main contemporary trait could be identified in a stronger role of interest-based politics, coupled with lower levels of civic culture.

Such brief outline of the main features of these four geographic areas sets the context for the comparison that follows. We ran separate analyses for the four different areas of the country, with interesting results. The parameter estimation is presented in Table 2.

**Tab. 2 - Parameter estimate for Italy 2001 (separate geographical zones) – ITANES 2001 data**

| Geographical zone        | Parameter         | Estimate | Std. Error | 95% Confidence Interval |             |
|--------------------------|-------------------|----------|------------|-------------------------|-------------|
|                          |                   |          |            | Lower bound             | Upper bound |
| Industrial Northwest     | B                 | .881     | .019       | .843                    | .919        |
|                          | Z                 | .000     | .639       | -1.264                  | 1.264       |
|                          | H                 | 10.951   | 3.149      | 4.730                   | 17.173      |
|                          | R-squared = 0.814 |          |            |                         |             |
| Former “White” Northeast | B                 | .935     | .027       | .881                    | .990        |
|                          | Z                 | .231     | 1.372      | -2.489                  | 2.950       |
|                          | H                 | 13.597   | 5.838      | 2.026                   | 25.169      |
|                          | R-squared = 0.750 |          |            |                         |             |
| “Red” Central Italy      | B                 | .855     | .016       | .824                    | .886        |
|                          | Z                 | 1.677    | .873       | -.049                   | 3.402       |
|                          | H                 | 7.318    | 2.771      | 1.837                   | 12.799      |
|                          | R-squared = 0.902 |          |            |                         |             |
| South                    | B                 | .899     | .017       | .865                    | .933        |
|                          | Z                 | 2.784    | 1.119      | .575                    | 4.993       |
|                          | H                 | 4.776    | 2.858      | -.867                   | 10.419      |
|                          | R-squared = 0.886 |          |            |                         |             |

The results show interesting differences among the various zones<sup>21</sup>. The first point worth mentioning is that *Z* both Northern regions is close to 0, meaning that, among least involved voters, the ideological dimension has almost no effect on the probability of voting for one of the two blocs: these voters seem to not distinguish (at least based on the L-R ideological dimension). Although, in the same regions there is a wide range of variations, since *Z+H* is 10,951 in the industrial Northwest and almost 14 in the former-White Northeast. This means that in both these regions highly involved voters are very clearly divided between the two blocs, with a very low permeability.

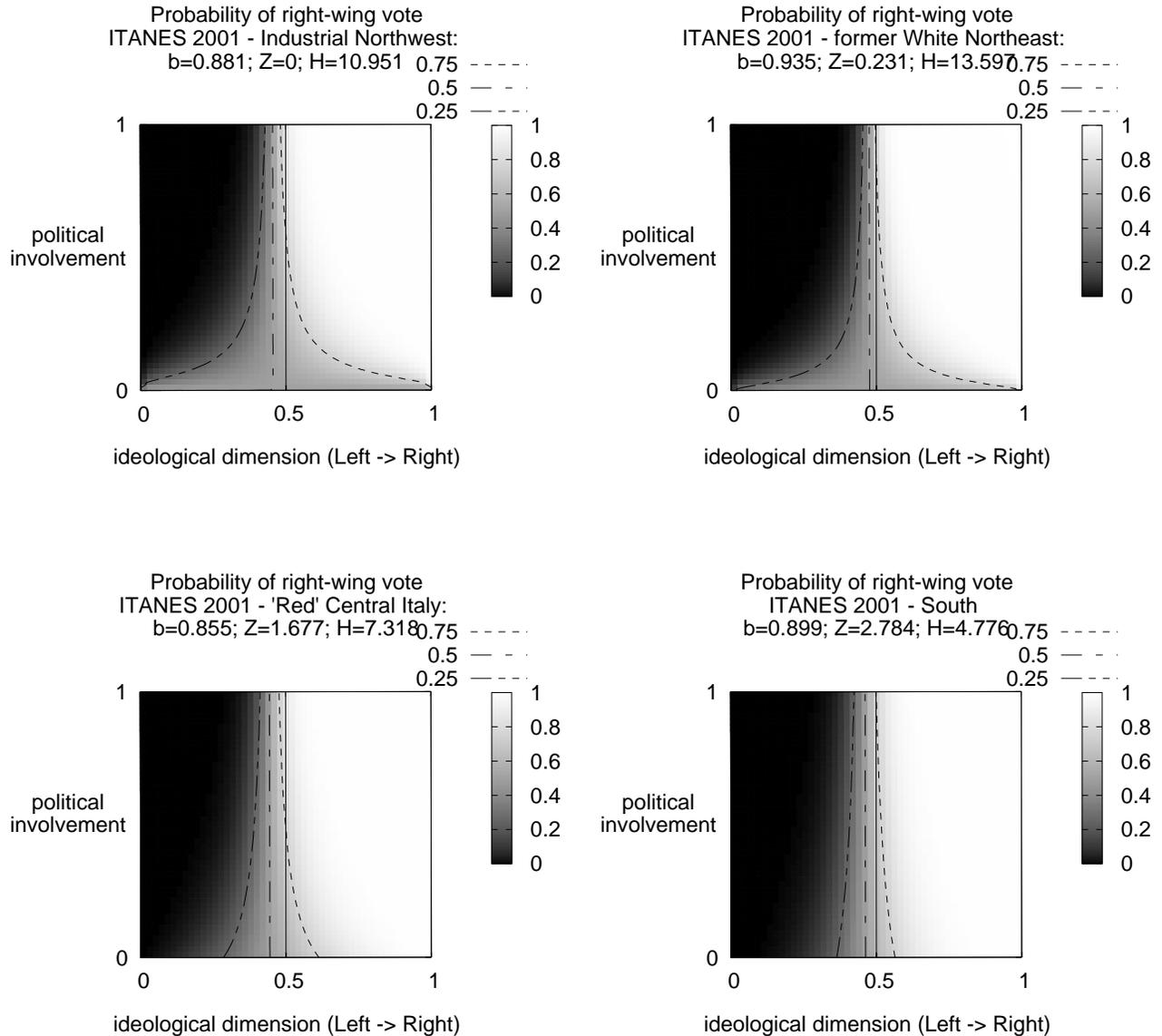
A somewhat different situation is seen in Central Italy. Here *Z* is 1.677, showing a degree of impermeability already among least involved voters. And, differences with the most involved are slightly less strong than in both Northern regions, with *Z+H* being about 9.

This effect is even stronger in the South. Here *Z* is highest: 2.784, meaning that even for least involved voters there is a strong relation between L-R position and bloc vote. *Z+H* is lowest, marking about 8: the situation among most involved voters is more similar to the least involved.

<sup>21</sup> General lower values of R-squared are due to the fact that, when disaggregating by geographical zone, grid cells are less populated, thus more prone to random error.

Finally, there seem not to be really important variations in bias across zones. It is, however, worth mentioning that, although the Red zone always has strong left-wing majorities, it still shows the overall national bias favouring the right-wing bloc. The results are briefly shown in Figure 7.

**Fig. 7 - L-R self-positioning and right-wing vote probability (shades of gray) – Estimated models for different geographical zones**



## Conclusions

A first set of conclusions is of methodological scope. By applying simple logical considerations and continuity, we were able to translate basic, primarily qualitative hypotheses into a logical quantitative model. Such a model makes precise quantitative predictions, and thus is quite easily

falsifiable. However, we found that the model fits the data well, even still when considering that it is a probabilistic model, only accounting for “spatial regularities” in voting behaviour<sup>22</sup>. The presence of adjustable parameters lets us explore the conformance of data to different aspects of our theoretical model, rather than allowing us to improve the fit with parameters that lack substantial meaning, and cannot be compared across different analyses. As a result, all the parameters estimated assume values that have a clear theoretical interpretation, and that point to different aspects of the theory. The mathematical form of the model already conveys a wealth of quantitative information, with data relevant in determining the exact parameters,

Going to the substantive results, our basic hypotheses seem to be in large part confirmed. Political involvement seems to matter for party competition. There is evidence of a higher permeability between party blocs among the voters who are least involved, meaning that for party competition this set of voters could prove a more rewarding target than in areas where more interested and involved (and educated) voters are to be found. The model also allowed us to trace differences among different geographical regions, pointing out that both Northern areas present strong differences between the extremes of the involvement dimension, with a highly polarized situation among the most involved and an almost “flat” discernment level among the least, while the “Red zone” and the South show much more homogeneity. This result could be tentatively explained by the much higher levels of party attachment still registered in “Red” Central Italy, but such an explanation seems not plausible for the South.

Regarding the amount of bias observed, the case we examined shows that in 2001 the *Casa della Libertà* was able to exert an overall higher attraction on *all* voters, regardless of their spatial position. This caused a slight shift to the left of the “indifference line” where voters have equal probabilities of voting for either of the two blocs. We could not actually expect a very strong effect, for two basic reasons. The first is that the difference in votes between the two blocs was minimal in terms of actual election results: so we should not expect a strong difference in the “attraction force” exerted by the two blocs<sup>23</sup>. The second reason is that it is likely that a relevant part of the attraction exerted by the two blocs has also an effect on the ideological self-positioning of the respondent, thus making such effect undetectable. However, the effect is significant and of a plausible amount, since a slight degree of bias interpretively corresponds to the slight actual difference in votes.

Further tests are needed: in particular, evaluating whether similar patterns occur in other countries, or reversed patterns that still highlight the role of involvement<sup>24</sup>; and whether bias really reflects which of the two blocs is able to exert a higher attraction on all voters that is independent of ideology, and is able to affect who will win a close election. Still our preliminary tests of our quantitative logical model yielded results that are encouraging and which strongly suggest the model’s promise.

---

<sup>22</sup> The bare mathematical form derived only from logical considerations seems to already fit the data well. We tried tests for the whole country with the basic form of the model, that is, *without any parameter*, and it already has an R-squared of 0.517.

<sup>23</sup> There is no simple mapping from the bias estimated to the difference in actual votes, since three more variables should be taken into account: the representativeness of the sample; the spatial distribution of the actual voting population; and variable turnout. The only reasonable mapping is interpretive.

<sup>24</sup> See note 1.

## References

- Adams, J., Merrill, S. and Bernard Grofman. 2005. *A unified theory of party competition: a cross-national analysis integrating spatial and behavioral factors*. Cambridge: Cambridge University Press.
- Alvarez, R. Michael. 1997. *Information and elections, Michigan studies in political analysis*. Ann Arbor: University of Michigan Press.
- Black, D. 1958. *The Theory of Committees and Elections*. Cambridge: Cambridge University Press.
- Cartocci, R. 1987. "Otto risposte a un problema: la divisione dell'Italia in zone politicamente omogenee." *Polis* I(3): 481-514.
- Conway, M. Margaret. 2000. *Political participation in the United States*. 3rd ed. Washington, D.C.: CQ Press.
- De Sio, L. 2006 (forthcoming). "Declino o transizione? Il rapporto tra gli italiani e i partiti." In *I cambiamenti nella cultura politica degli italiani*, ed. Marco Maraffi. Bologna: Il Mulino.
- Delli Carpini, M. X. and Scott Keeter. 1996. *What Americans Know about Politics and Why it Matters*. New Haven: Yale University Press.
- Downs, A. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Galli, G. 1966. *Il bipartitismo imperfetto: comunisti e democristiani in Italia*. Bologna: Il Mulino.
- Grofman, B. 2004. "Downs and Two-party Convergence." *Annual Review of Political Science*, 7, pp. 25-46.
- Kendall, M. G., and A. Stuart. 1950. The Law of the Cubic Proportion in Election Results. *British Journal of Sociology* 1 (3):183-196.
- Macdonald, S. E., G. Rabinowitz, and O. Listhaug. 1995. Political Sophistication and Models of Issue Voting. *British Journal of Political Science* 25:453-483.
- Putnam, R. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press.
- Rivers, D. 1988. Heterogeneity in Models of Electoral Choice. *American Journal of Political Science* 32 (3):737-757.
- Taagepera, R. 2005a. "Beyond Regression: The Need for Logical Models" presented at the Third Conference of the Belgian Political Science Association -French community, Liège, 29-30 April 2005.
- 2005b. "Predictive vs. Postdictive Models." presented at the Third ECPR Conference, Budapest, 8-10 September 2005.
- Tufte, E. R. 1973. Relationship between Seats and Votes in 2-Party Systems. *American Political Science Review* 67 (2):540-554.
- Van Deth, J.W. and Elff, M. 2000. "Political Involvement and Apathy in Europe 1973-1998". *Arbeitspapiere - Mannheimer Zentrum für Europäische Sozialforschung* 33.
- Zaller, J. 1992. *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.

## Appendix: Proofs

Proof 1:  $\frac{y}{1-y} = \frac{x}{1-x} \Leftrightarrow y = x$

$$\begin{aligned}\frac{y}{1-y} &= \frac{x}{1-x} \\ \Rightarrow y(1-x) &= x(1-y) \\ \Rightarrow y - xy &= x - xy \\ \Rightarrow y &= x\end{aligned}$$

Proof 2:  $\frac{y}{1-y} = \left(\frac{x^b}{1-x^b}\right)^n \Leftrightarrow y = \frac{(x^b)^n}{(x^b)^n + (1-x^b)^n}$

$$\begin{aligned}\frac{y}{1-y} &= \left(\frac{x^b}{1-x^b}\right)^n \\ \Rightarrow y &= \frac{(x^b)^n(1-y)}{(1-x^b)^n} \\ \Rightarrow y &= \frac{(x^b)^n}{(1-x^b)^n} - y \frac{(x^b)^n}{(1-x^b)^n} \\ \Rightarrow y \left(1 + \frac{(x^b)^n}{(1-x^b)^n}\right) &= \frac{(x^b)^n}{(1-x^b)^n} \\ \Rightarrow y &= \frac{\frac{(x^b)^n}{(1-x^b)^n}}{\left(1 + \frac{(x^b)^n}{(1-x^b)^n}\right)} \\ \Rightarrow y &= \frac{(x^b)^n}{(x^b)^n + (1-x^b)^n}\end{aligned}$$