Monitoring the 1982 Spanish Socialist Victory: A Bayesian Analysis

JOSÉ M. BERNARDO*

An estimate of how people intended to vote was obtained four weeks before the Spanish general elections of October 1982. The evolution of opinion was followed during the campaign. Using a small sample of polling stations, the victory of the Spanish Socialist Party was predicted with great accuracy only two hours after the polls closed. A Bayesian hierarchical model was used.

KEY WORDS: Bayesian logistic regression; Hierarchical models; Information measures; Opinion polls.

1. INTRODUCTION

The general elections held in Spain on October 28, 1982, gave the power to the Spanish Socialist Party, in an impressive landslide victory, for the first time since the Civil War.

This article describes the following methods used to predict the outcome of the election: (a) monitoring how people intended to vote throughout the electoral campaign and (b) producing early predictions of the final results, only two hours after the polls closed, preceding by several hours any other accurate description of the election outcome.

The argument lies entirely within the Bayesian framework and presents three aspects that are novel compared with published analyses of electoral results. Specifically:

1. Rather than using randomly selected polling stations, the team used a logarithmic information measure to identify a number of polling stations that could be considered representative of the political behavior of the population under study and took random samples within them.
2. To distribute the undecided vote obtained in the opinion polls among all parties taking part in the election, the team discarded ad hoc procedures in favor of a Bayesian logistic regression analysis.
3. To make inferences about the proportion of votes obtained by each party, and hence about their expected number of seats, from either the results of the opinion polls or the early electoral returns, we made a Bayesian analysis of an appropriately chosen hierarchical model.

This permitted a probability distribution over the possible configurations of the Parliament to be obtained, an impossible task from a classical viewpoint.

In Section 2, I outline the Spanish electoral system. Section 3 discusses the problem of data selection. In Section 4, I present the model used, and the Bayesian analysis for it is described in Section 5. Section 6 deals with the classification of undecided votes. In Section 7, I illustrate the method with the results obtained in the province of Valencia, and I conclude with some final remarks in Section 8.

2. THE SPANISH ELECTORAL SYSTEM

The Spanish Parliament consists of two Houses. The lower house, the Congreso de los Diputados, consists of 350 seats; the leader of the party or coalition that has a plurality of the seats is appointed by the King to be President of the Government. The upper house has no role in this context.

The country is divided into 52 electoral units, or provincias (provinces), each of which elects a number of diputados (members of the lower house) that is roughly proportional to its population, with a correction to enhance the representation of the less populated areas. Thus Valencia, a comparatively heavily populated provincia (with 6% of the country's population) elects 15 diputados rather than 20, which would be its proportional share.

The seats in each province are divided among the parties that obtain at least 3% of the vote in the province, according to a corrected proportional system usually known as the d'Hondt rule. (This is a requirement of the electoral law negotiated by all political parties after the death of the dictator.) The Jefferson–d'Hondt rule, invented by Thomas Jefferson nearly a century before Victor d'Hondt rediscovered and popularized the system (Balinski and Young 1980, p. 18), is also used in Argentina, Austria, Belgium, Finland, Switzerland, and West Germany; a variation is used in Denmark, Norway, and Sweden.

According to the d'Hondt rule, to distribute s seats among the, say, m parties that have overcome the 3% barrier, one (a) computes the $m \times s$ matrix of quotients with general elements

$$z_{ij} = n_i/j, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, s,$$

* José M. Bernardo is Professor and Head, Department of Biostatistics, Facultad de Medicina, Ave. Blasco Ibañez 17, Valencia–10, Spain. The implementation of the methods described was the work of a team that included C. Armero, M.J. Bayarri, J. Ferrandiz, M. Rabena, L. Sanjuan, and M. Sendra. The author acknowledges illuminating discussions with Phil Brown in the early development of this model and helpful comments by the editor, associate editor, and a referee.

© Journal of the American Statistical Association
September 1984, Volume 79, Number 387
Applications Section

where n_i is the number of valid votes obtained by the ith party; (b) selects the s largest elements; and (c) allocates to party i a number of seats equal to the number of these s largest elements found in its corresponding row. Clearly, to apply the d'Hondt rule, one may equivalently use the proportion of valid votes obtained by each party, rather than the absolute number of votes.

Thus if, for example, 15 seats are to be distributed among four parties A, B, C, and D, which have obtained 53%, 29%, 6%, and 4%, respectively, of the valid votes (the rest being distributed among parties with less than 3% of the vote), one forms the matrix

	1	2	3	4	5	6	7	8	9	10	11
A	53	26.5	17.7	13.3	10.6	8.8	7.6	6.6	**5.9**	5.3	—
B	29	14.5	9.7	7.3	**5.8**	4.8	—				
C	**6**	3	—								
D	4	2	—								

from which A, B, C, and D would obtain 9, 5, 1, and 0 seats, respectively.

It is not difficult to verify that the d'Hondt rule provides a corrected proportional system that enhances the representation of the big parties to the detriment of the smaller ones, the correction being larger when the number of electoral districts increases. Indeed, in the preceding example, a proportional representation would yield 8.64, 4.73, .98, and .65, which are not far from the 9, 5, 1, 0 provided by the d'Hondt rule. A party such as D, however, with similar results in all provinces, would not have representation in the Parliament despite its 4% of the vote, which would imply, proportionally, 14 of the 350 seats.

Each electoral unit or province is divided into a variable number of polling stations, each containing between 500 and 2,000 electors living in a small, roughly homogeneous geographical area. An important feature of the Spanish electoral system is that the votes are counted in public in each polling station just after its closing time (8 p.m.). This means that at about 9 p.m., someone attending the counting may communicate by telephone to the analysis center the returns of about the first 200 scrutinized votes. These data are used to make early predictions of the results.

The mechanics of seat allocation imply that any sensible statistical analysis of Spanish electoral data has to proceed province by province, only combing in a final step the predictions of the number of seats to be won by each of the competing parties in each of the 52 provinces. Indeed, because of important regional differences deeply rooted in history, electoral data in a given region are only mildly relevant to a different one; for instance, the electoral results in Catalonia or the Basque Country are vastly different from those of Andalusia or Valencia. Consequently this model focuses on obtaining electoral predictions for a single province, although I shall later comment on the results of predicting for the provinces in combination. For purposes of illustration I shall use the 1982 electoral data from the province of Valencia, which elects 15 seats.

3. DATA SELECTION

Each province is divided into a large number of polling stations, averaging about 1,000 electors each. The province of Valencia is divided into 1,774 polling stations, for which the team of analysts had the electoral results, by polling station, obtained in the March 1979 general elections.

I wanted to identify which stations returned more representative results, in the sense of yielding estimates of the percentages of votes for each party that would be similar to those finally obtained for the whole area. This would eliminate possibly atypical areas and would reinforce the plausibility of the assumptions of a simple model. Thus I needed an appropriate distance among the corresponding probability distributions.

A number of related arguments suggested the use of the directed divergence. Indeed, a sensible measure (Kullback 1968, p. 6; Bernardo 1979a) of the loss to be expected if one uses an approximate distribution $Q = \{q_1, \ldots, q_n\}$, $q_i \geq 0$, $\sum q_i = 1$, rather than the true distribution $P = \{p_1, \ldots, p_n\}$, $p_i \geq 0$, $\sum p_i = 1$, to predict the proportion p_i of the vote that the ith party will obtain is given by

$$L(Q \mid P) = \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i},$$

which is easily shown to be nonnegative. A more sophisticated loss function could be defined in terms of the differences in parliamentary seats provided by P and Q through the d'Hondt rule.

For each of the polling stations of a given province where m parties with electoral possibilities took part, we traced the corresponding 1979 results $\{n_{i0}, n_{i1}, \ldots, n_{im}\}$, where n_{ij} is the number of 1979 votes of party j in polling station i and n_{i0} the aggregate number of votes obtained by small parties. By the argument given in Section 4, a Bayesian estimate of the probabilitiy θ_{ij} that an elector with characteristics similar to those living in the area covered by polling station i will vote for party j is

$$\hat{\theta}_{ij} = (n_{ij} + 1/2)/(N_i + (m + 1)/2), \quad j = 0, 1, \ldots, m,$$

where

$$N_i = \sum_{j=0}^{m} n_{ij}.$$

For each of the polling stations, we then computed

$$l_i = \sum_{j=0}^{m} \theta_j \log (\theta_j/\hat{\theta}_{ij}),$$

where θ_j was the proportion of votes obtained in the province by party j in the 1979 general elections. The smaller l_i was, the more representative the polling station i would be, according to this particular criterion. We assumed that this representativity would remain essentially unchanged and, accordingly, limited sampling to the areas covered by the more representative polling stations.

Some computer simulations using the 1979 results suggested that random samples of 50 electors within each of the 20 more representative polling stations in each province would produce predictions sufficiently accurate for purposes of this study.

If the 1982 results have a linear regression on the 1979 results, then stations should be chosen with a large variance for the 1979 results to ensure a good estimate of the slope of 1982 on 1979, which is precisely the opposite of the design we used. Empirical evidence on political data strongly suggest, however, that a single regression function of the results of one election on those of a previous one cannot be expected to hold over the entire population; on the contrary, the evolution in voting patterns may be expected to be different in sociologically diverse areas. On the other hand, the same empirical evidence also suggests that representative areas remain representative, thus allowing accurate predictions based solely on them.

4. THE MODEL

Let us consider a particular province in which m parties compete for s seats. Let θ_{ij} be the (unknown) probability that a citizen from the area covered by polling station i will vote for party j. Naturally these probabilities will change from one area to another, but, using the data selection procedure outlined above, those changes should be dramatic.

Let us assume that a random sample of size N_i is taken among citizens of polling station i, and let n_{ij} be the number who vote for party j, so $N_i = n_{i0} + n_{i1} + \cdots + n_{im}$, where n_{i0} is the number of votes for small parties (not considered in the analysis). The data may be either (a) the result of an opinion poll of size N_i in the area covered by polling station i, in which the vote intention is determined directly or indirectly for all the N_i citizens approached, or (b) the returns of the first N_i votes scrutinized in polling station i on election night.

I shall assume that $(n_{i0}, \ldots, n_{im})$ is a random sample from a multinomial distribution with parameters $\theta_{i0}, \ldots, \theta_{im}$, $\theta_{ij} \geq 0$, $\sum_j \theta_{ij} = 1$. From a Bayesian viewpoint, the information provided by $(n_{i0}, \ldots, n_{im})$ about the θ_{ij}'s is encapsulated in the corresponding posterior distribution of the θ_{ij}'s. Prior information on the θ_{ij}'s is difficult to specify because of (a) the large number of variables involved and (b) the complicated structure relating the θ_{ij}'s to one another. A Dirichlet prior is mathematically convenient and not difficult to assess (Bunn 1978), but it cannot usually accommodate (b); a feasible alternative may be to assume a logistic normal multivariate distribution (Aitchison and Shen 1980) and an appropriate hyperprior on its parameters. In our role as consultants, however, we were specifically asked by the politicians to avoid the use of prior subjective information and to report only conclusions that could be derived from the data. Thus we used the standard reference uninformative prior $\pi(\theta) \propto \prod \theta_{ij}^{-1/2}$, $i \in$ {polling stations}, $j = 1, \ldots, m$ (Jeffreys 1967, p. 184; DeGroot 1970, p. 174; Bernardo 1979b). The corresponding reference posterior distribution is Dirichlet with parameters $(n_{ij} + 1/2, j = 1, \ldots, m)$; hence the expected value of θ_{ij} given the data is $\hat{\theta}_{ij} = (n_{ij} + 1/2)/(N_i + m/2)$, and the corresponding log odds would be $\log\{\hat{\theta}_{ij}/(1 - \hat{\theta}_{ij})\}$.

This motivates the following definition and assumptions.

Definition 1. The posterior log odds x_{ij} that correspond to a vector $(n_{i0}, \ldots, n_{im})$ are defined to be

$$x_{ij} = \log(\hat{\theta}_{ij}/(1 - \hat{\theta}_{ij})) = \log((n_{ij} + \tfrac{1}{2})/(N_i - n_{ij} + (m - 1)/2)), \quad j = 1, \ldots, m, \quad (1)$$

where $N_i = n_{i0} + \cdots + n_{im}$. Note that we only need to define the log odds for the m parties actually considered in the analysis; the proportion θ_{i0} of votes for the small parties may be finally estimated, if desired, from the relation $\sum \theta_{ij} = 1$. Clearly, in a log odds scale, x_{ij} describes the sample strength of party j in polling station i.

Assumption 1. The vectors $x_i = (x_{i1}, \ldots, x_{im})$ of posterior log odds are approximately normally distributed within area i around a vector of means μ_i (which depends on the area), with a common precision matrix H_0; that is,

$$p(x_i \mid \mu_i, H_0) \simeq N_m(x_i \mid \mu_i, H_0). \quad (2)$$

I am intuitively assuming that (a) the x_{ij}'s may be seen as measures of political strength with an approximate normal distribution and (b) the statistical consequences of the relative political positions of the different parties, as described by the precision matrix, are constant throughout the electoral province.

This assumption will be only approximately true if all of the N_i are equal and if for all i, $\theta_{ij} \simeq \theta_j$ (i.e., if the sample size chosen within each polling station was the same and if the vote distribution within each polling station was similar). Indeed, this was the case with our final design, where a random sample of fixed size was taken in each polling station and the polling stations themselves were chosen (as described in Section 3) so that their vote distributions would be roughly similar.

Assumption 2. The vectors $\mu_i = (\mu_{i1}, \ldots, \mu_{im})$ of means are approximately normally distributed within the province; that is,

$$p(\mu_i \mid \delta, H_1) \simeq N(\mu_i \mid \delta, H_1), \quad (3)$$

where δ describes the electoral behavior of the province in log odds.

I am intuitively assuming that the variations in the political behavior of similar areas within a given province are the consequence of a large number of small, independent causes so that the central limit theorem may be invoked.

Assumptions 1 and 2 together produce a hierarchical model (cf. Lindley and Smith 1972) and may easily be combined. Indeed, if t is the total number of polling stations considered,

$$p(x_1, \ldots, x_t \mid \delta, H_0, H)$$
$$= \iint \cdots \int \prod_{i=1}^{t} N_m(x_i \mid \mu_i, H_0) \times N_m(\mu_i \mid \delta, H_1)\, d\mu_1 \cdots d\mu_m$$
$$= \prod_{i=1}^{t} \int N_m(x_i \mid \mu_i, H_0) \times N_m(\mu_i \mid \delta, H_1)\, d\mu_i$$
$$= \prod_{i=1}^{t} N_m(x_i \mid \delta, H_0(H_0 + H_1)^{-1} H_1), \quad (4)$$

so the vectors $x_1, \ldots, x_t$ of sample log odds may be seen as a random sample from an m-variate normal distribution centered in a vector $\delta = \{\delta_1, \ldots, \delta_m\}$ that describes the global electoral behavior of the province and has an unknown precision matrix $H = H_0(H_0 + H_1)^{-1}H_1$.

5. BAYESIAN ANALYSIS

The proposed model contains a vector of parameters of interest δ, the province log odds, and a matrix of nuisance parameters H. From a Bayesian point of view, one has to determine the posterior distribution of the parameter of interest δ. With the appropriate reference uninformative prior $\pi(\delta, H) \propto |H|^{-(m+1)/2}$ for δ and H, the reference posterior distribution of δ is

$$p(\delta \mid \text{data}) \propto |S + (\bar{x} - \delta)(\bar{x} - \delta)'|^{-t/2} \quad (5)$$

(Geisser and Cornfield 1963; Bernardo 1979b)—that is, an m-variate t distribution with mean $\bar{x}$, dispersion matrix $S/(t - m)$, and $t - m$ degrees of freedom—where

$$\bar{x}_j = \sum_{i=1}^{t} x_{ij}/t, \quad \bar{x} = (\bar{x}_1, \ldots, \bar{x}_m)$$
$$S = \sum_{i=1}^{t} \{(x_i - \bar{x})(x_i - \bar{x})'\}/t, \quad S = \{s_{jk}\}.$$

Posterior probability intervals for the components of δ are readily obtained from the marginal distributions based on (5). Specifically,

$$\Pr[\delta_j \in \bar{x}_j \pm h_\alpha \sqrt{\{s_{jj}/(t - m)\}}] = 1 - \alpha, \quad (6)$$

where h_α is the $1 - \alpha/2$ quantile of a normalized Student's t with $t - m$ degrees of freedom.

Since the δ_j's are the province log odds, an appropriate estimate for the proportion of votes ψ_j to be obtained by party j in the province will be obtained using the corresponding inverse transformation, so

$$\hat{\psi}_j = e^{\bar{x}_j}/(1 + e^{\bar{x}_j}), \quad j = 1, \ldots, m. \quad (7)$$

Similarly, probability intervals for the ψ_j's will be given by

$$e^{\alpha_j}/(1 + e^{\alpha_j}), \quad e^{\beta_j}/(1 + e^{\beta_j}), \quad j = 1, \ldots, m, \quad (8)$$

where (α_j, β_j) are the extremes of the probability intervals for the δ_j's defined in (6).

Important as they are, inferences about the proportion of votes are not the final object of the analysis; we are interested instead in a probability distribution over the different possible configurations of the province's seats.

Let F be the function that associates a configuration of seats to each distribution of votes by using the d'Hondt rule over those parties overcoming the 3% barrier. Thus

$$F\{(\psi_1, \ldots, \psi_m)\} = (s_1, \ldots, s_m), \qquad s_i \geq 0, \sum s_i = s, \quad (9)$$

means that were the m parties to obtain proportions of votes given by $\psi_1, \ldots, \psi_m$ in the province analyzed, then they would get, respectively, $s_1, \ldots, s_m$ of the s seats corresponding to that province.

Then the probability of a given configuration $(s_1, \ldots, s_m)$ will be given by

$$\Pr\{s_1, \ldots, s_m\}$$
$$= \iint \cdots \int_{R_1} p(\psi_1, \ldots, \psi_m \mid \text{data}) d\psi_1, \ldots, d\psi_m$$
$$= \iint \cdots \int_{R_2} p(\delta_1, \ldots, \delta_m \mid \text{data})\, d\delta_1, \ldots, d\delta_m, \quad (10)$$

where $R_1 = \{\psi_1, \ldots, \psi_m), (\psi_1, \ldots, \psi_m) \in F^{-1}(s_1, \ldots, s_m)\}$, R_2 is the image of R_1 by $\delta_j = \log\{\psi_j/(1 - \psi_j)\}$, and $p(\delta_1, \ldots, \delta_m \mid \text{data})$ is given by (5). The multiple integral in (10) does not have an analytic solution and was calculated in simulation. Thus 1,000 vectors $(\delta_1, \ldots, \delta_m)$ were generated according to (5), transformed to $(\psi_1, \ldots, \psi_m)$ by $\psi_j = e^{\delta}j/(1 + e^{\delta}j)$, and further transformed to $(s_1, \ldots, s_m)$ by (9); the probability of each configuration of seats was then estimated as its corresponding relative frequency.

Once the probability distribution over the possible configurations is obtained using (10), it is easy, and informative, to compute the marginal distributions over the number of seats to be won by each party.

6. UNDECIDED VOTE: PROBABILISTIC CLASSIFICATION

If the model described is to be used to analyze an opinion poll, it is necessary to know how every citizen in the sample intends to vote. Only a proportion (about 80%) of the people asked made clear how they intended to vote, so the problem arises of how to classify the undecided vote.

Naturally, the questionnaire included not only intended vote but also other indicators such as age, sex, educational level, profession, leadership preferences, political position in a general spectrum, vote in past elections, and so on. As a consequence, we had (a) a data bank that consisted of the indicators and the preferred party of a number of citizens and (b) the indicators of a (smaller) number of undecided voters. We then used a Bayesian probabilistic classification procedure (Bernardo 1983;

Bernardo and Bermudez 1984), based on the complete data, to determine the probability distribution of the vote of each undecided citizen, given his or her associated attributes. We assumed that citizens with the same profile, as defined by their indicators, would have the same political behavior, as described by their probabilities of voting for the various parties. This seems a reasonable assumption, provided sufficiently rich profiles are used.

If a particular undecided person with indicators $y_1, \ldots, y_k$ provided the probability distribution over the parties $\{p_1, \ldots, p_m\}$, where $p_j = \Pr\{\text{voting party } j \mid y_1, \ldots, y_k\}$, then we would add p_j ($j = 1, \ldots, m$) to the number of citizens expressing their intention of vote for party j. This procedure not only provides a sensible, complete (probabilistic) classification of the undecided vote, which takes into account the obvious fact that people with similar profiles (i.e., similar indicators) should have similar probability distributions, but it also produces interesting information about the voting tendencies of segments of the population.

7. EXAMPLE: PROVINCE OF VALENCIA

The team of analysts conducted two opinion polls based on domiciliary visits to 50 citizens randomly chosen from the electoral census in each of the 20 most representative areas of the province.

The polls were taken on October 8 and October 21, 1982; 77% of the people interviewed in the first poll and 81% in the second expressed how they intended to vote. The remaining, still undecided citizens were probabilistically classified, using logistic regression on sex, political tendency (in a spectrum from 1, extreme left, to 9, extreme right), and vote in the 1979 election, because this was the combination of indicators found to minimize the expected loss of misclassification.

We found, for instance, that a man with political tendency 3 who voted socialist in 1979 had a probability .12 of voting PCE (communist), .47 of voting PSOE (socialist), .06 of voting UPV (nationalist), .05 of voting CDS (center–left), .05 of voting UCD (center, in government), and .08 of voting AP (conservative), the rest being distributed among smaller options.

On October 8 and October 21 (four weeks and one week, respectively, before election day), we provided accurate estimates of how people intended to vote at that time, thus producing useful political information within the campaign.

On election night, we used the returns of the first 100 valid votes in the 20 selected representative polling stations to produce results at 10 p.m. and the final results of the same polling stations to produce our definitive fore-

Table 2. Evolution Through the Electoral Campaign of the Probability Distribution Associated With the Allocation of the 15 Valencia Seats Among the Parties Considered

Probability of Configuration	Distribution by Party				
	PSOE	AP	UCD	CDS	PCE
Poll 1					
.27	8	3	2	1	1
.18	9	2	2	1	1
.16	9	3	1	1	1
Poll 2					
.40	9	3	1	1	1
.23	10	3	1	1	0
.13	9	4	1	1	0
Oct. 28, 1982					
.17	10	4	0	0	1
.58	*10*	*5*	*0*	*0*	*0*
.16	9	5	0	0	1
.03	9	6	0	0	0
.06	11	4	0	0	0
Oct. 29, 1982					
.74	*10*	*5*	*0*	*0*	*0*
.12	9	5	0	0	1
.12	9	6	0	0	0
.02	11	4	0	0	0

NOTE: Italicized entries are the final election results.

Table 1. Evolution Through the Electoral Campaign of the Percentages of Valid Votes Predicted for Each of the Parties Considered

	Party				
	PSOE	AP	UCD	CDS	PCE
Poll 1					
Prediction	43.9	15.7	10.2	5.6	6.9
90% CI	39.0, 48.9	12.6, 19.4	7.9, 13.2	4.0, 7.8	5.1, 9.2
Poll 2					
Prediction	50.7	18.4	7.7	9.2	5.2
90% CI	47.3, 54.2	13.3, 24.9	5.0, 11.8	7.0, 11.9	4.0, 6.7
Oct. 28, 1982					
Prediction	53.4	27.2	3.3	1.8	5.0
90% CI	49.0, 57.7	23.5, 31.2	2.3, 4.6	1.1, 2.9	4.0, 6.2
Oct. 29, 1982					
Prediction	53.8	29.5	3.9	2.2	4.6
90% CI	50.1, 56.8	26.6, 32.6	3.3, 4.6	1.8, 2.7	3.7, 5.6
Final Results	53.3	29.4	4.4	2.3	5.3

NOTE: 90% CI = 90% Bayesian confidence interval. See text for explanation of other abbreviations.

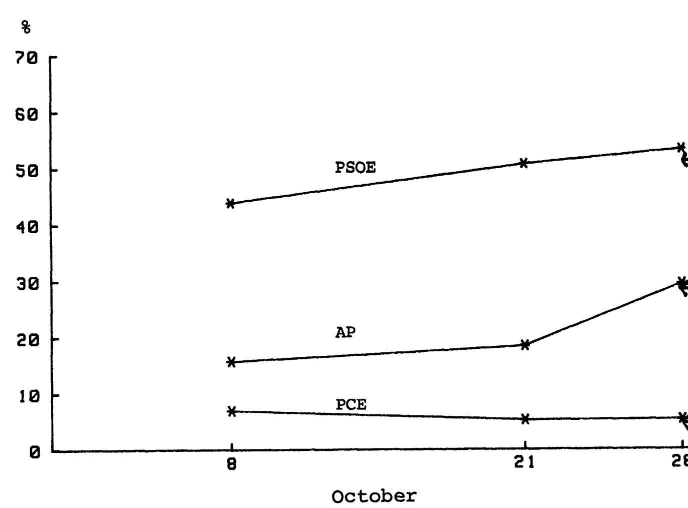

Figure 1. Evolution of the expected values as a function of time. The arrows indicate the final official results of the three main parties.

Table 3. Evolution Through the Electoral Campaign of the Probability Distribution of the Number of Seats in Valencia Obtained by Each Party

Party	Seats	Poll 1	Poll 2	Oct. 28	Oct. 29
PSOE	8	.44	.06	—	—
	9	.48	.62	.19	.24
	10	.04	.32	.74	.74
	11	—	—	.06	.02
AP	2	.24	.10	—	—
	3	.74	.68	—	—
	4	.02	.20	.23	.02
	5	—	—	.74	.87
	6	—	—	.03	.12
UCD	*0*	.01	.04	1.00	1.00
	1	.31	.88	—	—
	2	.68	.08	—	—
CDS	*0*	.11	—	1.00	1.00
	1	.89	.93	—	—
PCE	*0*	—	—	.67	.88
	1	1.00	1.00	.33	.12

NOTE: Italicized entries are the final election results.

casts at midnight, thus providing the media with accurate estimates of the results several hours before provisional results were given by anyone else.

The results obtained appear in the adjoining tables. Table 1 presents the percentages of votes successively expected for each of the five main political parties competing in Valencia and their .90 highest probability density (HPD) intervals. Figure 1 is a graph of the evolution of the voter's intention through the electoral compaign.

Table 2 gives the probabilities successively attached to the configurations of the 15 seats of Valencia that were compatible with the data, and Table 3, those associated with the number of seats that each party could obtain. The final official results are given in each table; by comparing the predictions with them, it can be verified that the method worked extremely well.

8. FINAL COMMENTS

Combining the probability distributions of seat configurations that correspond to each of the 52 provinces to plot a probability distribution over the possible configurations of the lower house is theoretically very easy, although computationally difficult because of the large number of operations involved. An efficient algorithm was derived, however, that makes use of the Monte Carlo simulations required by the numerical integrations of (10).

The final integrated results for the whole country, predicting an absolute majority for the socialists of about 201 seats of the 350 (the final official result, known one week later, was 202) was announced at midnight, hours before any approximate provisional result could be given by either the government or any other organization.

The model can certainly be improved. A hierarchical model directly defined over the multinomial probabilities may be more appropriate, if computationally more involved. Prior information about the relationships among different regions may be introduced by means of partial exchangeability assumptions in the second stage of the hierarchical model. A full decision-theoretical analysis should be performed to decide the optimal design of the sample, both in terms of area selection methods and in terms of sample size.

We believe, however, that the extremely good predictions obtained suffice to prove that the main novel points of the analysis—namely, the use of representative areas, the Bayesian analysis of appropriately chosen hierarchical models, and the probabilistic classification of the undecided vote in opinion polls—are certainly important steps in the right direction.

[*Received January 1983. Revised March 1984.*]

REFERENCES

AITCHISON, J., and SHEN, S.M. (1980), "Logistic Normal Distribution: Some Properties and Uses," *Biometrika*, 67, 262–272.

BALINSKI, M.L., and YOUNG, H.P. (1980), *Fair Representation*, New Haven, Conn.: Yale University Press.

BERNARDO, J.M. (1979a), "Expected Information as Expected Utility," *Annals of Statistics*, 7, 686–690.

——— (1979b), "Reference Posterior Distributions for Bayesian Inference" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 41, 113–147.

——— (1983), "Bayesian Logistic Diagnostic Distributions," Universidad de Valencia, unpublished manuscript.

BERNARDO, J.M., and BERMUDEZ, J.D. (1984), "The Choice of Variables in Probabilistic Classification," *Bayesian Statistics 2*, eds. J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith, Amsterdam: North-Holland.

BUNN, D.W. (1978), "The Estimation of a Dirichlet Prior Density," *Omega*, 6, 371–373.

DEGROOT, M.H. (1970), *Optimal Statistical Decisions*, New York: McGraw-Hill.

GEISSER, S., and CORNFIELD, J. (1963), "Posterior Distributions for Multivariate Normal Parameters," *Journal of the Royal Statistical Society*, Ser. B, 25, 368–376.

JEFFREYS, H. (1968), *Theory of Probability*, Oxford, England: Oxford University Press.

KULLBACK, S. (1968), *Information Theory and Statistics*, New York: Dover Press.

LINDLEY, D.V., and SMITH, A.F.M. (1972), "Bayes Estimates for the Linear Model," *Journal of the Royal Statistical Society*, Ser. B, 34, 1–42.