REVIEW OF SOCIOLOGY, 2012/4: 4–16.

The Black Day of Public Opinion Research Revisited

Bayesian reanalysis of the Hungarian opinion poll results from 2002

Gergő ZÁVECZ
zavecz_gergo@ceu-budapest.edu

ABSTRACT: This paper reanalyzes the forecasts of the 2002 Hungarian parliamentary elections based on a rarely used statistical approach, Bayesian analysis. In contrast to the frequentist approach of statistics, this method enables the researcher to include prior information in the analysis. It may be relevant in the case of the elections under investigation, since although many polling companies had measured a higher support for the Hungarian Socialist Party in the first three months of 2002, before the elections incorrectly they all predicted the victory of the Fidesz-Hungarian Democratic Forum coalition. The empirical analysis is based on two samples used by Ipsos and by Tárki to produce their forecasts published eight days before the elections. The results show that in this particular case the Bayesian analysis is more precise than the frequentist analysis of the same data or the forecasts the same opinion polling companies published.

KEYWORDS: 2002 Hungarian parliamentary elections, Bayesian statistics, election forecasts

Introduction

"The black day of public opinion research (Gallup)"; László Kovács: "It was not opinion polls that we wanted to win, but rather the elections (Szerető 2010)". These quotes reflect the general tone of commentaries made by experts and politicians on the performance of public opinion companies after they had failed to forecast the winner of the 2002 Hungarian parliamentary elections. In Table 1, the actual results (parties' vote shares on the party list in the first round), the forecasts opinion polling companies published eight days before the elections and Election Day forecasts are compared.

Table 1. The actual results at the Hungarian 2002 parliamentary elections and the forecasts of the opinion polling companies (numbers in parentheses indicate percentages)

Parties[1]	Actual results	Tárki		Medián		Ipsos		Gallup	
		I.	II.	I.	II.	I.	II.	I.	II.
Hungarian Socialist Party	42.05	37.8	36-40	40	42	38	38-42	37	38-39
Fidesz-Hungarian Democratic Forum	41.07	45.8	44-48	42	44	43	43-47	47	46-47
Alliance of Free Demo-crats	5.57	5.9	5-7	7	5	7	4-5	5	6-7
Hungarian Justice and Life Party	4.37	3.7	3-6	3	4	5	4-5	4	3-4
Centrum Party	3.9	3.7	3-4	4	2	3	3	4	3-4
Workers' Party	2.16	?	3-4	?	2	3	2	?	1-2
Independent Smallholders' Party	0.75	?	0-1	?	1	1	1	?	0.5-1

I. shows forecasts published eight days before the elections, II. shows Election Day forecasts
Sources: actual results: http://www.valasztas.hu/parval2002/so02/ered_ind.htm
forecasts published eight days before the elections: Karácsony – Lakatos (2006), 199.
Election Day forecasts: http://www.gallup.hu/Gallup/release/ppref020408.htm

What makes this case even more interesting is that most opinion polling companies measured a higher support for the Hungarian Socialist Party in the months before the elections. As it may be seen in Table 2, their predictions changed right before the election in March 2002 or in the forecast published eight days before the election. In Table 2 the published predictions by two of these companies (Ipsos and Tárki) for the two biggest parties are presented.

Table 2. Predictions of Ipsos and Tárki before the 2002 elections (numbers in parentheses indicate percentages)

Parties	January 2002		February 2002		March 2002		Eight days before		Election Day	
	Tárki	Ipsos	Tárki	Ipsos	Tárki	Ipsos	Tárki	Ipsos	Tárki	Ipsos
Hungarian Socialist Party	45	43.7	45	44.6	37	42.8	37.8	38	36-40	38-42
Fidesz-Hungarian Democratic Forum	43	38.1	42	36.6	47	41.4	45.8	43	44-48	43-47

Sources: Ipsos database; Tárki: http://www.tarki.hu/hu/research/elect/gppref_table_03.html

1 Only the seven biggest parties are presented and discussed here and later in the analysis since all the other parties received only 0.13% of the votes. In the case of the forecasts published eight days before the elections by Tárki, Medián and Gallup the predicted vote shares for the Workers' Party and the Independent Smallholders' Party were not found.

Several papers have been published on the possible reasons of this failure – two of them have been identified by most authors. First of all, forecasters may have miscalculated the party preferences among voters who intended to vote for sure but either did not want to reveal the party they favored or were not sure in their party choice (and thus, they underestimated the Hungarian Socialist Party's vote share). Secondly, they may have miscalculated the turnout (Gallup; Karácsony – Lakatos 2006; Kolosi – Tóth 2002).[2]

Forecasts made by the opinion polling companies are based on the analysis of their population samples. They make propositions or in other words they infer on the population based on their samples. In the case of election forecasts based on the vote shares present in the sample (especially among those respondents who claim that they are going to vote for sure) they make propositions on the possible vote shares in the population – with the indication of margins of errors. This is the frequentist analysis of the sample. However, many experts go further, and build intuitive models in which they complement this kind of analysis with their political knowledge and expert intuitions. They do so in order to estimate the party preferences of respondents who are not sure about going to the elections. Since 2002 these methods have been developed for the reasons discussed (Kolosi – Tóth 2002). In sum, there are two main ways of forecasting election results. One is based on the frequentist analysis of the sample representative of the given population. The other is the frequentist analysis of these samples complemented with the intuitions of experts.

In this paper my aim is to reanalyze the samples used for forecasting the 2002 Hungarian parliamentary elections using a different philosophical approach – the Bayesian one. Bayesian statistics differ from frequentist statistics in several ways. Here I summarize the differences most relevant for election forecasts. Most importantly, a Bayesian analysis incorporates more than the sample gathered at a given time point (in this case right before the elections). The researcher's subjective assumptions and personal information on the parameter (in this case on the actual election results) are also included in the analysis in the form of the so-called prior distribution. The subjective assumptions may be based on either previous results or on expert opinions. In many cases instead of using such informative prior distributions which are based on subjective assumptions, researchers use non-informative prior distributions, which lack these assumptions, in order to keep certain advantages of the Bayesian analysis. This kind of analysis contradicts the underlying idea of Bayesian statistics – to incorporate personal assumptions about the parameter. Besides the prior distributions, based on the sample gathered at a given time point the researcher also describes the so-called likelihood

2 Some alternative explanations have also been published. One claims that the voters had been frightened by a possible Fidesz-Hungarian Democratic Forum landslide victory based on previous election poll results and thus changed their party choices (Letenyei – Takács 2003). However, this explanation is less relevant for the purposes of this paper since the mechanisms described by its authors may lead to imprecise results even if the forecast would measure the actual results right before the election – the effect of forecasts on the actual election results is assumed by this mechanism.

– which means political parties' observed vote shares in the sample in the case of an election forecast. The multiplication of the prior distribution and the likelihood are proportional to the so-called posterior distribution. In the Bayesian analysis the posterior distribution is used in order to make inference on the value of the parameter – in this case on political parties' actual vote shares.

Formally, this may be described in the following way: $f(\theta|x) \propto f(\theta) * f(x|\theta)$, where $f(\theta|x)$ is the posterior distribution, $f(\theta)$ is the prior distribution, $f(x|\theta)$ is the likelihood, θ is the unknown parameter investigated by the researcher, and x are the data about the parameter in the sample (for the methods, see: Horváth 2001; Lynch 2007; O'Hagan 2004; 2008). Both the posterior distribution and the likelihood are conditional. For the former it means that the inference on the parameter is (partly) conditional on the data $(\theta|x)$. For the latter it means that "the probability distribution for the data is conditional on the parameter" $(x|\theta)$ (O'Hagan 2008: 86).

The main point of doing a Bayesian analysis is that the researcher can incorporate his/her subjective assumptions and prior knowledge on the parameter under investigation. I hypothesize that it has a high relevance in the case of the 2002 Hungarian parliamentary elections. As it was seen in Table 2, opinion polling companies measured a higher support for the Hungarian Socialist Party in the months preceding the election. The Bayesian analysis helps the researcher to take into account these trends in the form of a prior distribution.[3] Thus, the results derived from the Bayesian reanalysis of the samples gathered before the 2002 elections may be closer to the actual results than the frequentist analysis of the same samples or the results provided by experts using their intuitive models building on both the samples and their political knowledge. This is the research hypothesis I am testing in this paper

The Bayesian approach for predicting election results

The use of the Bayesian approach in the retrospective reanalysis of samples used for forecasting elections is growing in the Western literature and is also spreading in predicting election results in advance (Kaplan – Barnett 2003). In this literature review I present papers which use polling results both in their prior distributions and in their likelihoods.

In these analyses the authors may use either non-informative or informative prior distributions. An example of the former is the work by Kaplan and Barnett (2003) in which they reanalyze samples used for forecasting the 1988 and 2000 US presidential elections. An example of the latter is the paper by Rigdon and his

3 The forecasts building on intuitive models also include subjective assumptions. However, the Bayesian approach is more formal and builds on methodological and statistical knowledge rather than on political expertise. The effects of prior information are more easily quantifiable in a Bayesian analysis than in an intuitive one done by an expert. Besides, as it is pointed out by Phillips and Edwards (1966) humans are conservative in a situation where they face new data but have prior information and subjective assumptions. In such a case, when a human calculates these distributions intuitively and not in a formalized way prior and posterior distributions are closer to each other than they should be according to Bayes' theorem. In other words, likelihood and thus the data in the sample are taken into account less than they should be.

co-authors (2009) in which they reanalyze the samples gathered before the 2004 US presidential elections using long-term electoral trends in the prior distribution. A similar work is done by Lynch (2007) in which he does the Bayesian analysis of a survey opinion poll done in Ohio just before the 2004 elections using the preceding three polls as prior knowledge for the prior distribution. He shows that while forecasts using the frequentist analysis of the sample gathered before the elections could not name the winner because the results were too close to call, the Bayesian analysis correctly would have predicted Bush's win.

These authors, most importantly Kaplan and Barnett (2003), point out two advantages of the Bayesian approach. First of all, if the results are too close to call (the difference between the vote shares of presidential candidates or parties is within the margin of error) in a forecast which is completely or partly based on the frequentist analysis of a sample, false predictions may occur. In a Bayesian analysis where prior information is also taken into account, especially when this prior information is the result of previous surveys, even if the previous survey results are too close to call, the accumulation of several too close to call survey results may show a pattern based on which the results of the election are no longer too close to call.

Secondly, with the use of the Bayesian approach it is statistically possible to make probability statements about the population parameters, while in the frequentist approach it is not meaningful to do so; Lynch (2007) makes such statements in his Bayesian analysis. In this case this means that in the Bayesian analysis one can attach probabilities to one party's win over the other for instance. It is the case since while the frequentist analysis can only deal with the uncertainty coming from random variability of the data in the sample based on which the inference on the population parameter is done, the Bayesian analysis claims that the population parameter is also uncertain and not fixed, thus allows to make subjective probability statements about it (O'Hagan 2004).

These arguments may provide further theoretical justification for my research, and based on them, it seems relevant to reanalyze the samples gathered before the 2002 Hungarian parliamentary elections with the Bayesian approach.

Research design

In the empirical analysis of my article one (actual results) plus two times four (various analyses of the samples gathered by two opinion polling companies) vote shares of political parties during the 2002 parliamentary elections are presented. The vote shares and the way they were computed are presented one by one.

Firstly, parties' actual vote shares are presented. I downloaded them from the following link: http://www.valasztas.hu/parval2002/so02/ered_ind.htm .

The second vote shares are from the forecast Ipsos published eight days before the elections. I collected them from the database gathered by Karácsony and Lakatos (2006). It is important to point out that this is an example of intuitive models which mix the frequentist analysis of the sample with experts' intuitions.

The third vote shares come from the frequentist analysis (without weights) of the sample used for the forecast Ipsos published eight days before the elections. The sample contains 3000 respondents out of which 1576 claimed that they were going to vote for sure and named the party they were going to vote for. In this analysis I gathered the vote shares based on respondents who were going to vote for sure. To get the vote shares I used the answers to the following question from the database: *For which party are you going to vote?*

Fourthly and fifthly, I reanalyzed the sample used for the forecast Ipsos published eight days before the elections with the Bayesian approach. Thus, the likelihood is based on the sample used for the forecast published eight days before the elections. To get the likelihood I used the answers to the same question presented in the case of the frequentist analysis and once again I focused only on the vote choices of respondents who claimed that they would attend the elections for sure. Regarding the prior distributions, in the fourth model, the prior distribution is non-informative which assumes uniform vote shares for all seven parties. This was done in order to make sure that I did not include subjective assumptions on the population parameter in the analysis. In the fifth model, the prior distribution is based on the samples of the polls Ipsos gathered in January, February and March 2002. It is to be noted that including three months in the prior distribution is proposed by Lynch (2007). The distribution of vote choices of respondents who claimed that they would attend the elections for sure was taken into account. The prior distribution is based on all these three samples. Technically, this means that the number of those who supported a given party in these three samples from three months was added up. I used the answers to the following question from the database to get the prior distribution: *If there were elections this Sunday for which party would you vote?* The question slightly differs from that used for the likelihood since here voting intentions about a hypothetical election were asked. In both the fourth and fifth models I multiplied the likelihood with the prior distributions – with the non-informative one in the former and with the informative one in the latter case. An important technical remark should also be made. Since the variables (about the respondents' vote choices) used in the analyses are measured on a multinomial scale, the likelihood is a multinomial distribution, and the priors are Dirichlet distributions. The Bayesian analysis was done in R with the MCMCpack package, using the MCmultinomdirichlet command.

The sixth, seventh, eighth and ninth vote shares are based on the analyses of the sample used for the forecast Tárki published eight days before the elections. The sixth vote shares are the ones Tárki actually published and are from the database by Karácsony and Lakatos (2006). The seventh vote shares come from my frequentist analysis (without weights) of the same sample. It contains 1519 respondents out of which 875 claimed that they were going to vote for sure and named the party they were going to vote for. I used the answers to the following question from the database to do the analysis: *If there were elections this Sunday for which party would you vote?*[4] In order to get the eighth vote shares I reanalyzed the sample with the Bayesian approach using a non-informative prior distribution and in order to get the ninth I did a Bayesian analysis using an informative prior distribution. In the latter case the prior distribution is based on the samples of the polls Tárki gathered in January, February and March 2002.[5]

In the next part I present and compare the nine vote shares.

Results

In Tables 3 and 4 parties' vote shares are compared with each other.

Table 3. Comparison of vote shares – Ipsos (numbers in parentheses indicate percentages)

Parties	Actual results	Ipsos published	Ipsos frequentist	Ipsos Bayesian – non-informative prior	Ipsos Bayesian – informative prior
Hungarian Socialist Party	42.05	38	38.8	38.8 (36.4 – 41.2)	41.7 (40.0 – 43.4)
Fidesz-Hungarian Democratic Forum	41.07	43	44.0	44.0 (41.5 – 46.4)	41.4 (39.7 – 43.1)
Alliance of Free Democrats	5.57	7	6.9	6.9 (5.7 – 8.2)	6.5 (5.7 – 7.4)
Hungarian Justice and Life Party	4.37	5	3.8	3.8 (2.9 – 4.8)	4.0 (3.4 – 4.7)
Centrum Party	3.9	3	2.6	2.6 (1.9 – 3.5)	2.4 (1.9 – 2.9)
Workers Party	2.16	3	2.6	2.6 (1.9 – 3.4)	2.1 (1.6 – 2.6)
Independent Small-golders' Party	0.75	1	1.1	1.2 (0.7 – 1.7)	1.2 (0.9 – 1.6)

Table 4. Comparison of vote shares – Tárki (numbers in parentheses indicate percentages)

Parties	**Actual results**	**Tárki published**	**Tárki frequentist**	**Tárki Bayesian – non-informative prior**	**Tárki Bayesian – informative prior**
Hungarian Socialist Party	42.05	37.8	38.5	38.5 (35.3 – 41.8)	41.3 (39.6 – 43.0)
Fidesz-Hungarian Democratic Forum	41.07	45.8	47.0	47.0 (43.5 – 50.2)	45.1 (43.4 – 46.9)
Alliance of Free Democrats	5.57	5.9	4.5	4.5 (3.2 – 5.9)	5.0 (4.3 – 5.8)
Hungarian Justice and Life Party	4.37	3.7	2.6	2.6 (1.7 – 3.8)	3.0 (2.5 – 3.7)
Centrum Party	3.9	3.7	3.3	3.3 (2.2 – 4.6)	2.4 (1.9 – 3.0)
Workers' Party	2.16	?	2.5	2.5 (1.6 – 3.7)	2.0 (1.5 – 2.5)
Independent Smallgolders' Party	0.75	?	0.7	0.7 (0.3 – 1.3)	0.8 (0.6 – 1.2)

4 This sample is regarded to be the forecast published eight days before the election by Karácsony and Lakatos (2006) as well. This is the second sample Tárki gathered in March 2002.

5 It is important to mention that the Centrum Party was in the Other Party category in January 2002. Based on the later trends almost everyone in the Other Party category was regarded as a voter for Centrum Party in this month.

Results for Ipsos show that the Bayesian analysis of the sample using informative prior distribution is the closest to the actual results in the case of the Hungarian Socialist Party, the Fidesz-Hungarian Democratic Forum coalition, the Alliance of Free Democrats, the Hungarian Justice and Life Party and the Worker's Party. Results for Tárki show that the Bayesian analysis with an informative prior distribution is most precise for the Hungarian Socialist Party, the Fidesz-Hungarian Democratic Forum coalition, the Worker's Party and the Independent Smallholder's Party.

The credible intervals derived from Bayesian analyses are indicated in the parentheses in the cells. Their interpretation would be that the population parameters (parties' actual vote shares) fall by 95% certainty in those intervals.[6] In the case of Ipsos all parties' actual vote shares fall in that interval except those of the Alliance of Free Democrats, the Centrum Party and the Independent Smallgolders' Party when an informative prior distribution is used, while only the actual vote shares gained by the Hungarian Justice and Life Party, the Workers' Party and the Independent Smallgolders' Party fall in that interval when the non-informative prior is included. For Tárki all parties' actual election results fall in that interval except those of the Fidesz-Hungarian Democratic Forum coalition, the Hungarian Justice and Life Party and the Centrum Party.

In order to evaluate and compare the accuracy of the different methods some measurements are needed (Campbell 2008: Table 1; Karácsony – Lakatos 2006: Table 1). One widely used method is the presentation of margin of errors (which is based on 95% confidence intervals in the case of frequentist analyses and on 95% credible intervals in the case of Bayesian analyses). In this paper I do not present the margin of errors since most authors presenting these numbers use a mistaken formula (Rudas 2006), a formula that should be used when the margin of error is computed for the whole sample and not only for the voters who are voting for sure. A modified margin of error should be used for the frequentist analysis but its computation is difficult, and is not meaningfully comparable with credible intervals derived from the Bayesian analysis, thus is not computed in this paper.

Thus, I compute one of the most widely used measurements for evaluating and comparing vote shares – mean absolute error from vote as proposed by Campbell (2008). For each party I compute the absolute difference between the actual result and the vote share provided by the given analysis and subsequently I add up these differences. Finally, the sum is divided by the number of parties which step would not be necessary since all the vote shares are related to one election with a given number of parties.[7]

6 Many authors incorrectly interpret confidence intervals in frequentist analyses in a very similar way. However, a 95% confidence interval means that with a huge number of repeated samples, 95% of the intervals computed from the samples would include the population parameter (Credible intervals, O'Hagan 2008).

7 Karácsony and Lakatos (2006) warn to be careful with this measurement since it treats a 2 percentage point difference as the same in case of both a party gathering a 1% vote share and a party getting 43% of the votes. However, since it is a widely used measurement (even Karácsony and Lakatos use its modified form in their paper), I also use it for comparison.

This measurement is computed for the Election Day forecasts Tárki, Medián, Ipsos and Gallup published, for the forecast Ipsos published eight days before the elections, for the frequentist and two Bayesian analyses of the sample used for the forecast Ipsos published eight days before the elections and for the frequentist and two Bayesian analyses of the sample used for the forecast Tárki published eight days before the elections. It would be meaningful to incorporate in the comparison the forecasts Tárki, Medián and Gallup published eight days before the elections. However, I have no access to the vote shares they predicted for the Workers' Party and the Centrum Party. It is also important to note that both intuitively and based on the results presented by Karácsony and Lakatos (2006), I expect a higher accuracy in the case of Election Day forecasts than in the case of the forecasts published eight days before the election – or than the reanalyses of the samples used for the latter forecasts.

Table 5. Comparison of the accuracy of forecasts

Forecast/analysis	Mean absolute error from vote
Tárki Election Day	1.65
Medián Election Day	0.89
Ipsos Election Day	1.21
Gallup Election Day	1.69
Ipsos published eight days before the elections	1.43
Frequentist reanalysis of the sample used for the forecast by Ipsos eight days before the elections	1.45
Bayesian reanalysis of the sample used for the forecast by Ipsos eight days before the elections using non-informative prior distribution	1.47
Bayesian reanalysis of the sample used for the forecast by Ipsos eight days before the elections using informative prior distribution	0.57
Frequentist reanalysis of the sample used for the forecast by Tárki eight days before the elections	1.90
Bayesian reanalysis of the sample used for the forecast by Tárki eight days before the elections using non-informative prior distribution	1.90
Bayesian reanalysis of the sample used for the forecast by Tárki eight days before the elections using informative prior distribution	1.20

Table 5 shows that the most accurate result is provided by the Bayesian reanalysis of the sample Ipsos used for the forecast published eight days before the elections using informative prior distribution where the prior distribution contains opinion polling results from January, February and March 2002. The vote shares provided by this analysis are more accurate than even the Election Day forecasts – which are surprising knowing the cited results by Karácsony and Lakatos (2006).

On the other hand, the Bayesian analysis using non-informative prior is not more accurate than the frequentist analysis of the same sample.

The Bayesian reanalysis of the sample Tárki used for the forecast published eight days before the election using informative prior distribution is not the most accurate forecast in general but is the most accurate one for that sample and for that company. The forecast Tárki published eight days before the elections is not included in Table 5 for reasons discussed above as their predictions for only five parties are available. It would not be meaningful to compare the mean absolute error from vote computed for seven parties with that computed for five parties, in this case the former would be probably lower. However, if one computes the mean absolute error from vote for these five parties, the following results may be concluded: it is 2.04 for the published results, 2.58 for the frequentist reanalysis of the sample Tárki used for the forecast published eight days before the election, 2.58 for the Bayesian reanalysis of the sample Tárki used for the forecast published eight days before the election using non-informative prior distribution and 1.64 for the Bayesian reanalysis of the sample Tárki used for the forecast published eight days before the elections using informative prior distribution. Based on this comparison, once again the Bayesian analysis with informative prior is the most accurate.

The Bayesian analysis also makes it possible to make probability statements about the population parameter, for instance about the chances of the Hungarian Socialist Party defeating the Fidesz-Hungarian Democratic Forum coalition. In the analysis of the Ipsos sample this chance is 2.9% when a non-informative prior and 56.0% when an informative prior is used. These percentages are 0.04% and 1% for Tárki. The second Bayesian analysis of the Ipsos sample yielded a higher probability of winning for MSZP.

In the conclusion I focus on the possible reasons of higher accuracy of the Bayesian analysis in this particular case.

Conclusion

Once again I would like to emphasize the theoretical arguments for the Bayesian approach. It makes it easier to incorporate prior information in a more formalized and quantifiable and probably less biased way than intuitive forecasting models. Besides, compared to analyses simply building on the frequentist analysis of samples, it enables us to take into account prior information and to make probability statements about the population parameter.

Besides, as it may be seen in the empirical part of the analysis in this particular case (the 2002 Hungarian parliamentary elections) the Bayesian analysis using informative prior distribution is more accurate than frequentist analyses or intuitive forecasting models for the samples Ipsos and Tárki provided. I do not intend to claim that the Bayesian analysis is superior, especially since it is only one Bayesian analysis in which there is a higher probability of the Hungarian Socialist Party victory. All that my results show is that in this particular case the Bayesian analysis seems to be more accurate for both opinion polling companies.

There might be several reasons behind this higher accuracy. As it has been noted before, opinion polling companies miscalculated the party preferences among voters who intended to vote for sure but either did not reveal the party they were going to vote for or were not sure in their party choice; they underestimated the support for the Hungarian Socialist Party within these groups. The Bayesian analysis corrected for this miscalculation by using prior distributions. In the surveys used for computing the prior distributions the Hungarian Socialist Party had a higher vote share than the Fidesz-Hungarian Democratic Forum coalition. Thus, the Bayesian analysis corrected the underestimation of the Hungarian Socialist Party supporters by including former trends. One might say that it is accidental that the Hungarian Socialist Party led in the months from which the prior information for the Bayesian analysis was collected and that the Bayesian analysis could not help in estimating the vote shares among those who did not reveal their party choices right before the elections. However, one may also argue that it is not accidental and it may be the case that the Hungarian Socialist Party supporters were showing their party preferences in a more manifest way a couple months before the elections and started to hide them only before the elections.

The question of accuracy of frequentist and Bayesian analyses cannot be decided based on one comparison. Therefore, in the future the reanalysis of other samples should be done. Besides, in the case of the 2014 Hungarian parliamentary elections the samples should be analyzed using all existing approaches. I hope that my article indicates the possibilities of the Bayesian analysis and points out that in the case of election forecasts even a mostly statistical approach can be as accurate as models built on years of political expertise.

Acknowledgements

I would like to thank Tamás Rudas for the help he gave me both in substantial and in methodological issues. Besides, I am also grateful to Juraj Medzihorsky, Eszter Timár and the two anonymous reviewers for all the useful advice they gave me. Last but not least, I would like to express my gratitude to the leaders of Ipsos Media, Opinion and Market Research Institute and to the TÁRKI Data Archive for the databases used for this analysis.

References

Campbell, J. E. (2008): Editor's Introduction: Forecasting the 2008 National Elections. *Political Science and Politics,* 41 (4): 679–682.

Credible Interval. Wikipedia. http://en.wikipedia.org/wiki/Credible_interval accessed on 24 July 2012.

Gallup*: A közvélemény-kutatás fekete napja.* http://www.gallup.hu/Gallup/release/ppref020408.htm accessed on 21 March 2012.

Horváth J. (2001): (A Bayes-statisztika és alkalmazása). In *BGF Tudományos Évkönyv*. Budapest: BGF, 218–225.

Kaplan, E. H. – Barnett, A. (2003): A New Approach to Estimating the Probability of Winning the Presidency. *Operations Research*, 51 (1): 32–40.

Karácsony G. – Lakatos Zs. (2006): Választási előrejelzések Magyarországon, 1990–2006. Sikerek, tévedések és tanulságok. In Angelusz R. – Tardos R. (Szerk.): *Mérésről mérésre*. Budapest: DKMKA, 185–220.

Kolosi T. – Tóth I. Gy. (2002): Egy tévedés története. In Kolosi, T. – Tóth, I. Gy. – Vukovich Gy. (Szerk): *Társadalmi riport 2002*. Budapest: Tárki, 339–367.

Letenyei L. – Takács K. (2003): Önpusztító jóslat – Miért veszítheti el a közvélemény-kutatások szerint népszerűbb párt a demokratikus választásokat? *Századvég*, 27: 83–99.

Lynch, S. M. (2007): *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. Berlin: Springer.

MCMCpack description for R. http://cran.r-project.org/web/packages/MCMCpack/MCMCpack.pdf accessed on 21 March 2012.

O'Hagan, A. (2004): Bayesian statistics: principle and benefits. In van Boekel, M.A.J.S. Stein, A. – van Bruggen, A. H. C. (eds): *Bayesian Statistics and Quality Modelling in the Agro-Food Production Chain*. Kluwer Academic, 31-45.

O'Hagan, A. (2008): The Bayesian Approach to Statistics. In Rudas, T. (ed): *Handbook of Probability: Theory and Applications*. Thousand Oaks: Sage, 85–100.

Phillips, L. D. – Edwards, W. (1966): Conservatism in a Simple Probability Inference Task. *Journal of Experimental Psychology,* 72 (3): 346–354.

Rigdon, S. E. et al. (2009): A Bayesian Prediction Model for the U.S. Presidential Election. *American Politics Research*, 37 (4): 700–724.

Rudas T. (2006): A hibahatár a becsült mennyiség függvényében – a mért pártpreferenciák téves értelmezésének egyik forrása. In Angelusz R. – Tardos R. (Szerk.): *Mérésről mérésre*. Budapest: DKMKA, 17–40.

Szerető Sz. (2010): *Pezsgő a hűtőben? Magyar Nemzet Online* http://mno.hu/vezercikk/pezsgo_a_hutoben-251500 accessed on 21 March 2012.

TÁRKI Omnibus 2002/1. *Social-political attitudes in Hungary among the adult population*. 2002-01-08 – 2002-01-15. Budapest: TÁRKI Data Archive.

TÁRKI Omnibus 2002/2. *Social-political attitudes in Hungary among the adult population.* 2002-02-05 – 2002-02-12. Budapest: TÁRKI Data Archive.

TÁRKI Omnibus 2002/3/A. *Social-political attitudes in Hungary among the adult population*. 2002-02-26 – 2002-03-06. Budapest: TÁRKI Data Archive.

TÁRKI Omnibus 2002/3/B. *Social-political attitudes in Hungary among the adult population*. 2002-03-19 – 2002-03-24. Budapest: TÁRKI Data Archive.