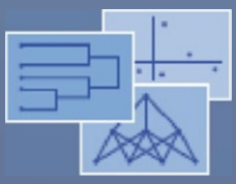
Studies in Classification, Data Analysis, and Knowledge Organization

Isabella Morlini
Tommaso Minerva
Maurizio Vichi *Editors*

Advances in Statistical Models for Data Analysis



Springer

Studies in Classification, Data Analysis, and Knowledge Organization

More information about this series at
http://www.springer.com/series/1564

Isabella Morlini • Tommaso Minerva •
Maurizio Vichi
Editors

Advances in Statistical Models for Data Analysis

Springer

Editors
Isabella Morlini
Department of Economics "Marco Biagi"
University of Modena & Reggio Emilia
Modena, Italy

Tommaso Minerva
Department of Communication and Economics
University of Modena & Reggio Emilia
Reggio Emilia, Italy

Maurizio Vichi
Department of Statistics
University of Rome "La Sapienza"
Roma, Italy

ISSN 1431-8814
Studies in Classification, Data Analysis, and Knowledge Organization
ISBN 978-3-319-17376-4 ISBN 978-3-319-17377-1 (eBook)
DOI 10.1007/978-3-319-17377-1

Library of Congress Control Number: 2015946232

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

This volume contains peer-reviewed selected contributions presented at the 9th biannual meeting of the Classification and Data Analysis Group (CLADAG) of the Italian Statistical Society that took place in Modena from September 18 to September 20, 2013. The conference brought together not only theoretical and applied statisticians working in Italy but also a number of specialists coming from nine different countries and was attended by more than 180 participants, including those who participated in a special session for young researchers. The conference encompassed 122 presentations organised into two plenary talks, two semi-plenary talks, 11 specialized sessions, 11 contributed sessions, eight coordinate sessions and a poster session. The main emphasis on the selection of the plenary and semi-plenary talks and on the call of papers was put on classification, data analysis and multivariate statistics, to fit the mission of CLADAG. However, many chosen contributions regarded related areas like machine learning, Markov models, structural equation models, statistical modelling in economics and finance, education and social sciences and environment. We would like to express our gratitude to all members of the Scientific Program and in particular to the Chair of the committee Francesco Palumbo. We also thank the local organizing committee, the session organizers, the invited speakers, the chairs and the discussants of all specialized sessions. We thank the authors of the contributions in this volume and the referees who spent time in carefully reviewing the papers and giving useful suggestions to the authors for improving theirs papers. We are largely indebted to the referees and to everyone who contributed their work to this volume. Finally, we thank Alice Blank from Springer for the cooperation provided in the publication of this volume.

Modena, Italy — Isabella Morlini
Modena, Italy — Tommaso Minerva
Roma, Italy — Maurizio Vichi
March 2015

tomavene@unict.it

Contents

Meta-Analysis of Poll Accuracy Measures: A Multilevel Approach

Rosario D'Agata and Venera Tomaselli

Abstract Following a meta-analysis approach as a special case of multilevel modelling, we identify potential sources of dissimilarities in accuracy measures of pre-election polls, carried out during Parliamentary elections in Italy from 2001 to 2008. The predictive accuracy measure, computed to compare the pre-electoral poll result to the actual result, is the dependent variable and the poll characteristics are the explanatory variables and are introduced in a hierarchical model. In the model each outcome is affected by a specific sampling error assumed to have a normal distribution and a known variance. The multilevel model approach decomposes variance components as well as meta-analysis random models. We propose a multilevel approach, in order to make the estimation procedure easier and more flexible than in a traditional meta-analysis approach.

Keywords Multilevel models • Pre-election polls

1 Meta-Analysis: A Case of Hierarchical Modelling

Envisaging to meta-analysis as a *special* case of multilevel modelling [2], we identify the potential sources of differences, by the estimation of an effect size over all the predictive accuracy measures of poll results. In this aim, we specify a random effects hierarchical model for a meta-analysis study of the measures, where each poll result is a level-1 unit and the poll is a level-2 unit [6], to check relationships between explanatory variables and dependent variables [3].

In order to compute an average effect size we can employ fixed or random effects models [1]. Specifying a fixed effects model, we assume that the true effect size is always the same across all studies. Formally [6]:

$$d_j = \delta_j + e_j \tag{1}$$

R. D'Agata (✉) • V. Tomaselli
University of Catania, Vitt. Emanuele II, 8 Catania, Italy
e-mail: rodagata@unict.it; tomavene@unict.it

I. Morlini et al. (eds.), *Advances in Statistical Models for Data Analysis*,
Studies in Classification, Data Analysis, and Knowledge Organization,
DOI 10.1007/978-3-319-17377-1_8

where:

– d_j is the outcome of study j ($j = 1, \ldots, J$)
– δ_j is the population value of the outcome of the j-th study
– e_j is the sampling error for this j-th study

As a consequence, the only error source is that produced by random sampling error or error e_j *within* studies, assumed to have normal distribution with known variance $\sigma^2_{e_j}$ [10].

Since we can suppose that the effect size across the studies will be similar but not identical, we can estimate different effect sizes across all studies. In this case, we specify a random effects model, in which the effect size, from each primary study, is estimated as a mean of a distribution of different true effect sizes δ_j across the studies, with an error term which is variance between studies. The observed effect d_j in (1) is sampled from a distribution of effects with true effect δ_j and variance $\sigma^2_{u_j}$ [7]. In turn, δ_j is a function of the mean of all true effects (γ_0) plus between studies error (u_j). Formally:

$$\delta_j = \gamma_0 + u_j \tag{2}$$

We can therefore rewrite (1) as:

$$d_j = \gamma_0 + u_j + e_j \tag{3}$$

– δ_j is the effect observed in the j-th study
– γ_0 is the estimate for the mean outcome across all the studies
– u_j is the between studies residual error term
– e_j is the within studies error term.

This is an *intercept only* or *empty* model, equivalent to the random effects model for meta-analysis [4], in which the variance of the residual errors $\sigma^2_{u_j}$ not equal to 0 and significant indicates that the outcomes across the studies are heterogeneous. In the model the effect size estimates are affected by two error sources: random sampling error or within studies error (e_j) and variance among the true effect sizes or between studies error (u_j).

In the multilevel approach to meta-analysis the dependent variable is the effect size of j-study. As data of primary level-1 units we use only summary statistics—i.e.: p-value, mean, correlation coefficient, odd-ratio, etc.—varying across studies as level-2 units [6].

Following the multilevel approach, in a random effects model, we can separate the variance of study outcomes into two components [5]:

– Within studies variance as sampling variance
– Between studies variance, due to the differences across the study results, computed in our application as predictive accuracy measures.

If the between variance is statistically significant, we can assess that the study outcomes are heterogeneous. To explain such a heterogeneity we include in a random effects model the study characteristics as explanatory predictors of the differences found in predictive accuracy measures across the studies. Estimating the effect size in the case of heterogeneous expected results by means of multilevel modelling is simpler than by traditional meta-analysis methods, because a multilevel approach is more flexible [6]. Furthermore, we can avoid the clustering of studies due to heterogeneous effect sizes across the studies. So, we do not need to identify any variable defining the membership of studies in a cluster.

Employing a multilevel approach (2) can be written as follows:

$$\delta_j = \gamma_0 + \gamma_1 Z_{1j} + \gamma_2 Z_{2j} + \cdots + \gamma_k Z_{kj} + u_j \tag{4}$$

where:

- δ_j is the effect size assumed as varying across the studies
- γ_0 is the mean of all true effects
- Z_{kj} are covariates as explanatory variables (study features)
- γ_k are the coefficients
- u_j is the error term, representing the differences across the studies, assumed to have normal distribution with known variance $\sigma^2_{u_j}$.

By substituting Eq. (4) into Eq. (1), the model can be written as:

$$d_j = \gamma_0 + \gamma_1 Z_{1j} + \gamma_2 Z_{2j} + \cdots + \gamma_k Z_{kj} + u_j + e_j. \tag{5}$$

So, the effect size estimate δ_j depends on study features Z_{kj}, on the error term u_j and on sampling error of each study e_j. The variance of uj ($\sigma^2_{u_j}$) could be considered a level-2 variance and indicates how much the outcomes vary across the studies.

In order to specify a multilevel model explaining the variance between studies, firstly we estimate an *empty* model (3) to check the homogeneity of outcomes, testing a null-hypothesis where the variance of the residual errors $\sigma^2_{u_j}$ is equal to 0 [8]. In meta-analysis with small sample and small variances, the Wald z-test is inaccurate for testing the variances [6]. Moreover, it is based on the assumption of normality and the variances have a χ^2-distribution with $df = j - k - 1$, where j is the number of studies and k is the number of covariates introduced into the model. So, we have to compute the deviance difference χ^2-test on the variances based on the sum of the squared residuals divided by their sampling variances or standard errors [6]. Formally:

$$\chi^2 = \sum_j \left[\frac{d_j - \hat{d}_j}{s.e.\left(d_j\right)} \right]^2. \tag{6}$$

If the null hypothesis is rejected, we have to estimate the proportion of variance due to the study characteristics or variance between the level-2 units. So, as in traditional

multilevel modelling, we can compute the ρ intra-class correlation coefficient (ICC), as the ratio between the level-2 variance $\sigma^2_{u_j}$ and the total variance $(\sigma^2_{u_j} + \sigma^2_{e_j})$. In formula:

$$\rho = \frac{\sigma^2_{u_j}}{\sigma^2_{u_j} + \sigma^2_{e_j}}. \tag{7}$$

In multilevel meta-analysis we have only level-2 variables. We can therefore only calculate the level-2 variance. A reduction of ρ indicates how the considerable extent of the covariates (the features of the study) affects this variance [7].

2 The Accuracy Measure of Pre-Electoral Polls

Unlike other sample surveys, for pre-election polls, we can make comparisons between poll results and actual voting results. For pre-election poll data, we suppose that how poll respondents indicate they will vote and how they actually vote in a subsequent election will correspond. If a poll result reflects the same distribution of voting preferences as happens in the following election, we have an accurate predictor. The more a predictor of an election outcome is able to provide unbiased estimates of electoral preferences, the more accurate it is. In order to measure how accurate a poll outcome is, we choose to use a A_{ij} poll accuracy measure as a predictor of an election result. By transforming poll outcomes into accuracy measures, we standardize all results thus making them comparable to one another.

A_{ij} measure was used for the first time to assess the predictive ability of pre-election polls in the U.S. Presidential elections of 1948, 1996, 2000 and also in the 2002 election for the Offices of Governor and Senator [9]. A_{ij} measure[1] is computed as the ratio obtained by dividing two odds:

$$A_{ij} = \ln \left\{ \left[s_{ij} / \left(1 - s_{ij}\right) \right] / \left[S_j / \left(1 - S_j\right) \right] \right\} \tag{8}$$

where:

- s_{ij} is the proportion of respondents favoring the s-competitor (party, coalition or candidate) in the i-th poll referring to the j-th population;
- $1 - s_{ij}$ is the proportion of respondents favoring all other competitors in the same i-th poll, for the same j-th population;
- S_j is the real proportion of votes polled by the same S-competitor in the same j-th population;

[1] A_{ij} measure is not affected by the size of the undecided voter category. Furthermore, it is standardized for the real election result. So, it is possible to study by means of a meta-analysis approach the origin of bias of the polls across different elections for race and time.

If the between variance is statistically significant, we can assess that the study outcomes are heterogeneous. To explain such a heterogeneity we include in a random effects model the study characteristics as explanatory predictors of the differences found in predictive accuracy measures across the studies. Estimating the effect size in the case of heterogeneous expected results by means of multilevel modelling is simpler than by traditional meta-analysis methods, because a multilevel approach is more flexible [6]. Furthermore, we can avoid the clustering of studies due to heterogeneous effect sizes across the studies. So, we do not need to identify any variable defining the membership of studies in a cluster.

Employing a multilevel approach (2) can be written as follows:

$$\delta_j = \gamma_0 + \gamma_1 Z_{1j} + \gamma_2 Z_{2j} + \cdots + \gamma_k Z_{kj} + u_j \tag{4}$$

where:

- δ_j is the effect size assumed as varying across the studies
- γ_0 is the mean of all true effects
- Z_{kj} are covariates as explanatory variables (study features)
- γ_k are the coefficients
- u_j is the error term, representing the differences across the studies, assumed to have normal distribution with known variance $\sigma^2_{u_j}$.

By substituting Eq. (4) into Eq. (1), the model can be written as:

$$d_j = \gamma_0 + \gamma_1 Z_{1j} + \gamma_2 Z_{2j} + \cdots + \gamma_k Z_{kj} + u_j + e_j. \tag{5}$$

So, the effect size estimate δ_j depends on study features Z_{kj}, on the error term u_j and on sampling error of each study e_j. The variance of uj ($\sigma^2_{u_j}$) could be considered a level-2 variance and indicates how much the outcomes vary across the studies.

In order to specify a multilevel model explaining the variance between studies, firstly we estimate an *empty* model (3) to check the homogeneity of outcomes, testing a null-hypothesis where the variance of the residual errors $\sigma^2_{u_j}$ is equal to 0 [8]. In meta-analysis with small sample and small variances, the Wald z-test is inaccurate for testing the variances [6]. Moreover, it is based on the assumption of normality and the variances have a χ^2-distribution with $df = j - k - 1$, where j is the number of studies and k is the number of covariates introduced into the model. So, we have to compute the deviance difference χ^2-test on the variances based on the sum of the squared residuals divided by their sampling variances or standard errors [6]. Formally:

$$\chi^2 = \sum_j \left[\frac{d_j - \hat{d}_j}{s.e.\left(d_j\right)} \right]^2. \tag{6}$$

If the null hypothesis is rejected, we have to estimate the proportion of variance due to the study characteristics or variance between the level-2 units. So, as in traditional

multilevel modelling, we can compute the ρ intra-class correlation coefficient (ICC), as the ratio between the level-2 variance $\sigma^2_{u_j}$ and the total variance $(\sigma^2_{u_j} + \sigma^2_{e_j})$. In formula:

$$\rho = \frac{\sigma^2_{u_j}}{\sigma^2_{u_j} + \sigma^2_{e_j}}. \tag{7}$$

In multilevel meta-analysis we have only level-2 variables. We can therefore only calculate the level-2 variance. A reduction of ρ indicates how the considerable extent of the covariates (the features of the study) affects this variance [7].

2 The Accuracy Measure of Pre-Electoral Polls

Unlike other sample surveys, for pre-election polls, we can make comparisons between poll results and actual voting results. For pre-election poll data, we suppose that how poll respondents indicate they will vote and how they actually vote in a subsequent election will correspond. If a poll result reflects the same distribution of voting preferences as happens in the following election, we have an accurate predictor. The more a predictor of an election outcome is able to provide unbiased estimates of electoral preferences, the more accurate it is. In order to measure how accurate a poll outcome is, we choose to use a A_{ij} poll accuracy measure as a predictor of an election result. By transforming poll outcomes into accuracy measures, we standardize all results thus making them comparable to one another.

A_{ij} measure was used for the first time to assess the predictive ability of pre-election polls in the U.S. Presidential elections of 1948, 1996, 2000 and also in the 2002 election for the Offices of Governor and Senator [9]. A_{ij} measure[1] is computed as the ratio obtained by dividing two odds:

$$A_{ij} = \ln \left\{ \left[s_{ij} / \left(1 - s_{ij}\right) \right] / \left[S_j / \left(1 - S_j\right) \right] \right\} \tag{8}$$

where:

- s_{ij} is the proportion of respondents favoring the s-competitor (party, coalition or candidate) in the i-th poll referring to the j-th population;
- $1 - s_{ij}$ is the proportion of respondents favoring all other competitors in the same i-th poll, for the same j-th population;
- S_j is the real proportion of votes polled by the same S-competitor in the same j-th population;

[1] A_{ij} measure is not affected by the size of the undecided voter category. Furthermore, it is standardized for the real election result. So, it is possible to study by means of a meta-analysis approach the origin of bias of the polls across different elections for race and time.

– $1 - S_j$ the actual proportion of votes polled by all other competitors in the same j-th population.

By dividing the poll odds by the election odds, we can obtain the value of odds ratio and, thus, the value of the A_{ij} measure as the natural logarithm of the odds ratio between the number of respondents who declare their intention to vote for the s_{ij}-competitor and those who intend to vote for all the others $(1 - s_{ij})$ in the i-th poll and for the j-th population, and the real number of votes for each of the two groups (S_j and $1 - S_j$) in the following election. The transformation of odds ratio by calculating its natural log is used to create a symmetric measure around 0 and to simplify the computation of the variance [9], taking into account the sampling error of the poll measure, assuming normal distribution and with a known variance.

Let s_{ij} and $1 - s_{ij}$ be random variables, where s_{ij} is the proportion of respondents preferring the s_{ij}-competitor and $1 - s_{ij}$ is the proportion of respondents who do not prefer the s_{ij}-competitor, in the i-th poll referring to the j-th population with sample size n_{ij}, with $[s_{ij} + (1 - s_{ij})] = 1$. Let $p(s_{ij})$ be the probability of preferring the s_{ij}-competitor and $[1 - p(s_{ij})]$ be the probability of not preferring the s_{ij}-competitor. The covariance matrix (Cov) of the vector $[s_{ij}, (1 - s_{ij})]$ is:

$$Cov\begin{bmatrix} s_{ij} \\ 1 - s_{ij} \end{bmatrix} = \frac{1}{n_{ij}} \begin{bmatrix} p(s_{ij})\left[1 - p(s_{ij})\right] & -p(s_{ij})\left[1 - p(s_{ij})\right] \\ -p(s_{ij})\left[1 - p(s_{ij})\right] & p(s_{ij})\left[1 - p(s_{ij})\right] \end{bmatrix} \tag{9}$$

so that the relative covariance matrix ($RelCov$) is:

$$RelCov\begin{bmatrix} s_{ij} \\ 1 - s_{ij} \end{bmatrix} = \frac{1}{n_{ij}} \begin{bmatrix} \left[1 - p(s_{ij})\right]/p(s_{ij}) & -1 \\ -1 & p(s_{ij})/\left[1 - p(s_{ij})\right] \end{bmatrix}. \tag{10}$$

The variance of A_{ij} measure (8) for each i-poll is computed as:

$$Var\left(A_{ij}\right) = 1/\left[n_{ij}s_{ij}\left(1 - s_{ij}\right)\right]. \tag{11}$$

A_{ij} measure may take on positive, negative or null values. A positive value indicates an s_{ij} bias. If the A_{ij} measure value is negative, the poll is biased by an overestimated share of $(1 - s_{ij})$ compared to $(1 - S_j)$ election result. The A_{ij} measure is equal to 0, when the odds ratio is equal to 1. This last result occurs only if the poll result and the real voting result are exactly the same.

To explain the variance of the measure, we can use A_{ij} as y-dependent variable and the poll features as predictors. We can assess whether there are significant relationships between the ability of the poll to predict the election results and the characteristics of the poll [9], including referred territorial area, customer, sampling procedures, survey methods, time poll period, sample size, number of days from poll to election, vote gap between the two competitors, polling agency, type of election, etc.

3 Meta-Analysis of Pre-electoral Poll Accuracy

In this study, we propose a meta-analysis of poll predictive accuracy measures in order to analyse their variance in a dataset of 42 pre-election polls. These polls have been carried out before the fortnight press blackout, previous the National elections from 2001 to 2008, and published on the official website: www.sondaggielettorali.it.

In order to assess the accuracy of each poll, we compute the accuracy measure by employing the formula [8]. Figure 1 shows the distribution of accuracy measures in the 42 polls. On average, we note that the Centre-Right electoral outcome (*AccDx*) is basically underestimated (−0.0432), while the Centre-Left coalition performance (*AccSn*) is overestimated (0.0754). Over the period of the elections considered, an improvement occurs in the ability of polls to accurately forecast results. For the election in 2008, accuracy measures computed both for *Centre-Right* coalition and for *Centre-Left* coalition are very near to 0. This could be due to an improvement in the quality of the methods used to conduct the pre-election polls such as sampling techniques and survey methods.

In order to evaluate the relationship between poll characteristics and accuracy measure, a meta-analysis is conducted using a multilevel approach. As the first step, we specify an empty model with the aim of checking heterogeneity across the polls as level-2 units. As the dependent variable, we choose the accuracy measure for *Centre-Left* coalition computed as shown in (8). Formally the model is:

$$AccSn_j = \gamma_0 + u_j + e_j \tag{12}$$

where:

– $AccSn_j$ is the accuracy measure for *Centre-Left* coalition in the j-th poll;

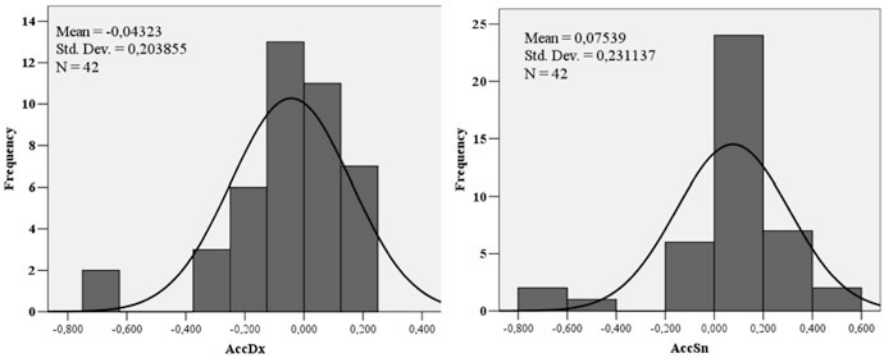


Fig. 1 Distribution of accuracy measures for the two coalitions

Table 1 Empty model: y = Accuracy of Centre-Left coalition ($AccSn_j$)

	β_{0j}	S.E.	Z	p-value
Fixed effects				
Intercept	0.078	0.035	2.229	<0.02
Random effects				
$\sigma^2_{u_j}$	0.047	0.011	4.273	<0.001
Deviance $= -4.839$				
Deviance difference test : $\chi^2 = 1089.7$; $df = 41$; p-value < 0.001				

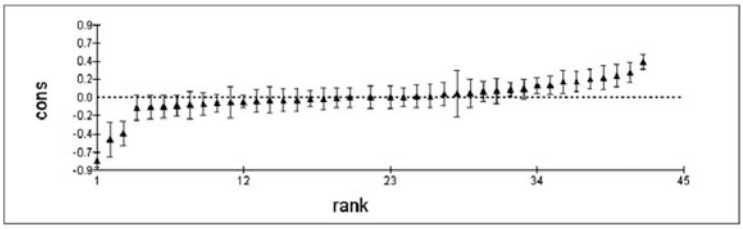


Fig. 2 Poll level residual plot (confidence intervals: 95 %) of empty model

– γ_0 is the estimate for the mean accuracy measure across all polls;
– u_j is the residual term for the j-th poll;
– e_j is the sampling error for the j-th poll computed by (11).

In the *empty* model[2] (Table 1), the value of the intercept (0.078) is significant (p-value $<$ 0.02) and confirms the overestimation of the Centre-Left result previously observed (Fig. 1). The random component ($\sigma^2_{u_j}$) indicates how much the accuracy measures vary across the polls. It is estimated as 0.047. In order to test the homogeneity of accuracy measures across the polls, we compute the deviance difference χ^2-test on the residuals to check for a null hypothesis where $\sigma^2_{u_j}$ is equal to 0. The test produces a χ^2 equal to 1089.7 (p-value < 0.001). So, we have to reject the null hypothesis, because the p-value indicates the presence of heterogeneity in accuracy measures across the polls.

In the plot of poll level residuals (Fig. 2), analysing the confidence intervals computed for the 42 polls, we note a group of about 12 polls where the confidence intervals for their residuals do not overlap 0. We can, therefore, observe 12 polls that differ significantly from the mean accuracy measures at the 5 % confidence level. Furthermore, the proportion of systematic variance, computed by means of an ICC (7), is 0.90 and informs us that in 90 % of polls the difference in accuracy measures is due to the features of the polls. The presence of heterogeneity and the value of the ICC allow us to continue the analysis with the aim of explaining in a

[2]The models are estimated by employing RML algorithm implemented in *MlWin* software.

Table 2 Complete model: y = Accuracy Centre-Left coalition ($AccSn_j$)

	β_{0i}	S.E.	Z	p-value
Fixed effects				
Intercept	0.675	0.224	3.013	<0.002
Customer: *agency* (Ref. Mass Media)	−0.179	0.056	−3.196	<0.001
Customer: *political organ* (Ref. Mass Media)	−0.280	0.110	−2.545	<0.006
Survey method: *CATI* e *CAWI* (Ref. *CATI*)	0.200	0.068	2.941	<0.002
Survey method: *CAWI* (Ref. *CATI*)	0.280	0.071	3.944	<0.001
Survey method: *CASI* (Ref. *CATI*)	0.369	0.14	2.636	<0.005
Sample size: $\ln(n_i/N_j)$	0.033	0.013	2.538	<0.006
Poll period (days)	0.055	0.021	2.619	<0.005
Days from poll to election	−0.019	0.006	−3.167	<0.001
Predicted gap	0.024	0.006	4.000	<0.001
Year: 2001 (Ref. 2008)	0.185	0.054	3.425	<0.001
Year: 2006 (Ref. 2008)	0.348	0.099	3.515	<0.001
Electoral winner: Centre-Left (Ref. Centre-Right)	−0.217	0.055	−3.945	<0.001
Random effects				
$\sigma^2_{u_j}$	0.007	0.002	3.500	<0.001
Deviance = −68.327				
Deviance difference test : $\chi^2 = 165.73$; $df = 29$; p-value < 0.001				

complete model the variance between polls in accuracy measures by means of the features of polls.

Table 2 shows that all predictors appear to be significant. Analysing the values of β_{0_j} coefficients, when the customer is a political organ or agency, the value of predictive accuracy measures tends to decrease compared to when the customer is one of the mass media (−0.179). The data collected by CAWI, either alone (0.280) or combined with *CATI* (0.200), appear to be linked to an increase in the value of accuracy. This is also true for *CASI* (0.369). All of the survey methods introduced in the model, compared to the *CATI* method, appear to have a positive relation to accuracy. The greater the sample size (0.033) and the longer the survey period (0.055), the more accurate the poll. The fewer days from poll to election (−0.019) and the greater the predicted gap (0.024) between the two coalitions, the more accurate the forecast is. Finally, if the winner is *Centre-Left* (−0.217), the accuracy measure decreases more than in elections won by the *Center-Right* coalition. Moreover, in the complete model the variance between the polls $\sigma^2_{u_j}$ is reduced from 0.047, observed in the empty model, to 0.007. The value of the deviance difference test is notably reduced from 1089.7 in empty model to 165.73, and it remains statistically significant, too. In addition, the proportion of systematic variance is reduced from 0.90 to 0.57. Comparing the poll residuals of the complete model, plotted in Fig. 3, with the poll residuals of the empty model (Fig. 2), we note that only one interval does not overlap 0.

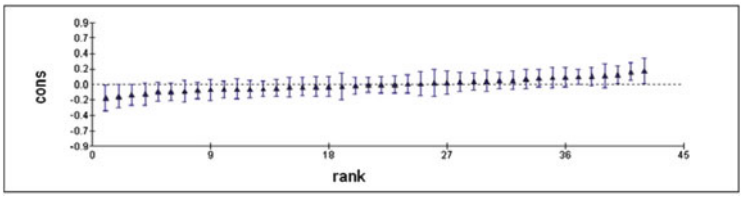


Fig. 3 Poll level residuals plot (confidence interval: 95 %) of complete model

4 Conclusions

The use of accuracy measures has made it possible to detect the predictive ability of each poll with a single value using a multilevel approach to meta-analysis. Specifying a random intercept model has allowed us to estimate the random component of the variance between the polls and to test the significance of heterogeneity among the results by means of the χ^2 residuals test. Thus, we were able to calculate the ICC in order to estimate the proportion of total variance, due either to the variance between the polls or the predictive accuracy measures across all the polls.

By means of a complete hierarchical model we obtained a reduction of the random component of the variance between the polls. So, the heterogeneity in the accuracy measures is explained by poll features as predictors in the estimated model. Nevertheless, a proportion of unexplained variance remains, as we note, both in the significance of the residuals test and the value of the ICC.

References

1. Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R.: Introduction to Meta-Analysis. Wiley, Chichester (2011)
2. Bryk, A., Raudenbush, S.W.: Hierarchical Linear Models: Applications and Data Analysis Methods. Sage, Newbury Park (1992)
3. Ellis, P.D.: The essential guide to effect sizes: statistical power, meta-analysis, and the interpretation of research results. Cambridge University Press, Cambridge (2010)
4. Hedges, L.V., Olkin, I., Statistiker, M., Olkin, I., Olkin, I.: Statistical Methods for Meta-Analysis. Academic, New York (1985)
5. Hox, J.: Multilevel Analysis: Techniques and Applications. Routledge, London (2010)
6. Hox, J., de Leeuw, E.: Multilevel models for meta-analysis. In: Reise, S.P., Duan, N. (eds.) Multilevel Modeling: Methodological Advances, Issues, and Applications, pp. 90–111. Lawrence Erlbaum Associates Publishers, Mahwah (2003)
7. Hunter, J.E., Schmidt, F.L.: Methods of meta-analysis: correcting error and bias in research findings. Sage, Thousand Oaks (2004)
8. Koetse, M.J., Florax, R.J., de Groot, H.L.: Consequences of effect size heterogeneity for meta-analysis: a Monte Carlo study. Stat. Methods Appl. **19**(2), 217–236 (2010)
9. Martin, E.A., Traugott, M.W., Kennedy, C.: A review and proposal for a new measure of poll accuracy. Public Opin. Q. **69**(3), 342–369 (2005)
10. Raudenbush, S.W., Bryk, A.: Hierarchical linear models: applications and data analysis methods, vol. 1. Sage, Thousand Oaks (2002)

DICHIARAZIONE SOSTITUTIVA DI ATTO DI NOTORIETA'
(sull'attribuzione della responsabilità dei singoli autori di lavori congiunti)
(Artt. 19 e 47 del D.P.R. 28.12.2000, n. 445)

La sottoscritta TOMASELLI Venera nata a Catania l'1/9/1961, residente a Pedara (provincia di CT) Corso Ara di Giove n. 12, C.A.P 95030, consapevole che, ai sensi dell'art. 76 del D.P.R. 445/2000, dichiarazioni mendaci, formazione o uso di atti falsi sono puniti ai sensi del codice penale e delle leggi speciali in materia,

DICHIARA

che nel lavoro a firma congiunta:

D'Agata R., TOMASELLI V. (2015). Meta-Analysis of Poll Accuracy Measures: A Multilevel Approach. In: Morlini I., Minerva T. and Vichi M. (Eds), Advances in Statistical Models for Data Analysis. (series: Studies in Classification, Data Analysis, and Knowledge Organization), p. 63-71, Berlin Heidelberg: Springer-Verlag. ISBN: 978-3-319-17377-1.

Peer-review.

il contributo degli autori è da considerarsi paritetico sotto ogni aspetto *e l'ordine degli autori è esclusivamente alfabetico.*

L'attribuzione della redazione dei paragrafi, tuttavia, è da intendersi nel seguente modo:

D'Agata R.: paragrafo 3
TOMASELLI V.: paragrafi 1, 2 e 4

La sottoscritta dichiara di essere informata, ai sensi dell'art. 10 della legge 675/96, che i dati sopra riportati saranno utilizzati nell'ambito del procedimento per il quale la presente dichiarazione viene resa.

Catania, 21/11/2016

La sottoscritta
Venera Tomaselli
Venera Tomaselli