Quick counts in the Mexican presidential elections: A Bayesian approach

Manuel Mendoza, Luis E. Nieto-Barajas[*]

Department of Statistics, ITAM, Mexico

ARTICLE INFO

ABSTRACT

In all democracies, anticipating the final results of a national election the same day the voters go to the polling stations is a matter of interest, for television stations and some civil rights organizations, for example. The most reliable option is a quick count, a statistical procedure that consists in selecting a random sample of polling stations and analysing their final counts to forecast the election results. In Mexico, a particularly important quick count is organized by the electoral authority. The importance of its results requires this exercise to be designed and executed with specially high standards far beyond those used in commercial studies of this type. In this paper, the model and the Bayesian analysis of the quick counts conducted by the Mexican authority, during the presidential elections in 2006 and 2012, are discussed.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

As in many other countries, in Mexico different organizations produce exit polls and quick counts the day of the election to get estimations of the final result. The largest and more sophisticated exercises are often those produced by the television networks although there are also studies generated by other media and, as usual in rather new democracies, a number of civil organizations, both foreign and local, report their results as well.

Predicting the result of an election has been of great interest for different authors. There are two main approaches. One based on opinion/exit polls before or during an election, and the other based on actual cast votes. When an opinion poll is conducted before the election day, a random sample of citizens who express their intention to vote, is selected. Once they reveal which party are willing to vote, this information is used to produce inferences. Alternatively, an exit poll is based on a random sample of polling stations. During the election day, a sample (usually of systematic type) of the citizens who cast their vote on each one of these stations is interviewed to know which party they voted for.

The use of opinion/exit polls has been severely criticised because they sometimes lead to misleading results. See for example, Brown et al. (1999) for the 1992 British general election, and Barreto et al. (2006) for the 2000 US presidential election. For the experience in Britain in 2005, Curtice and Firth (2008) not only discuss the difficulties that made impossible to conduct a traditional exit poll but describe a method which combines information from two consecutive elections. To this end, a panel of polling stations is defined where the information regarding each election is recorded by an exit poll. These data are used to estimate the change in each party's share of the vote. These estimates and the final results for the first election are used to forecast the results of the election in course. Inferences are successful although the method relies on specific characteristics of the British electoral system. Additionally, Anand and Jenkins (2004) raise some concerns about the fairness of opinion and exit polls in India since they claim can influence voting behaviour.

For the second approach based on final counts data, one possibility is to use the flow of results as they are produced when the polling stations close. This information is not the result of a random selection although a statistical model can be used to relate these final counts with those of previous elections in the same stations. In this regard, Bernardo and Girón (1992) proposed a Bayesian hierarchical model to predict the unobserved swings (difference between the votes from present and past elections). They work at polling station level and assume the swings to be exchangeable. On the other hand, Pavia-Miralles (2005) uses multivariate regression models to predict the current election data based on several past

---

* Corresponding author. Department of Statistics, ITAM, Rio Hondo 1, Col. Progreso Tizapan, 01080, Mexico D.F., Mexico.
E-mail addresses: mendoza@itam.mx (M. Mendoza), lnieto@itam.mx (L.E. Nieto-Barajas).

elections at the poll level and carries out a frequentist analysis. In these two cases, the authors acknowledge that the final counts they use to fit their models, are not random. They treat the information as random and claim that their estimates improve as more polls become available.

Working with the first available poll-by-poll final results (not at random), has the risk to produce misleading results, specially in tight elections where the competing candidates votes are very close to each other. There are some strategies to make a non random sample to be representative. For instance, Sedransk and Clyde (1966) propose a post stratification technique based on population sizes and past elections. Interestingly, Brown et al. (1999) describes a method which combines nonrandom final counts data with prior information arising from an exit poll. Other recent example which uses non random final counts data is Fisher (2015).

Final counts data can also be obtained from a random sample of polling stations. Once these stations close and produce final counts, the information is analysed to obtain forecasts. This is the type of exercise usually known as quick count and is used by a number of organisations to anticipate the final results of an election. In particular, a quick count is considered a useful tool against fraud and other illegal practices (Estok et al., 2002). Statistical methods for the production of estimates in quick counts, usually rely on the well-known techniques for inference in survey sampling studies (e.g. Cochran, 2001). The basic problem is then to estimate both the total number of valid votes and the total number of votes in favour of each candidate. The proportion of interest is then estimated using the corresponding ratio. No explicit model is used to describe the raw data (nonparametric assumption) although a normal model is used to approximate the sampling distribution of the estimates based on asymptotic theory.

From a Bayesian point of view, one contribution that deals with actual cast votes is that of Bernardo (1984). He uses an information measure to define a set of polling stations whose final results were close to the national ones in a previous election (non random selection). For each of these stations, he collects the first 100 cast votes and uses this partial counts to produce inferences. Bernardo assumes a two stages model with a multinomial distribution for the votes in each polling station. The posterior mean of the proportion of votes in the station in favour of each party is obtained as an estimate. Finally, the distribution of the vector of log-odds of these estimated proportions is approximated with a multivariate normal and a hierarchical prior depending on the true national proportions is used to get the posterior distribution of interest. This type of study is not a quick count since the selected poll stations are not a random sample and only partial counts data are used from each station.

Basing the predictions on past election results might be technically correct and useful but is politically unacceptable for a quick count organized by the electoral authority in Mexico. Parties could incorrectly claim that the resulting forecasts are deliberately biased in favour of the *status quo*. Interestingly, political parties never questioned the sample design used in the Mexican quick counts which makes use of past electoral information.

In this article we propose a Bayesian parametric model which relies on final counts data collected from the polling stations in a random sample. This model was used by the National Electoral Institute (INE), formerly Federal Electoral Institute (IFE), during the 2006 and 2012 presidential elections in Mexico. We illustrate the performance of our model with the 2006 election which is the tightest election ever organised in Mexico. Results show that our model is reliable providing good forecasts with the right precision to call the winner.

The use of Bayesian models in political science is becoming a common practice. For instance, Darmofal (2009) compares the performance of Bayesian spatial frailty models versus non spatial and non frailty models using the U.S. House members' position announcements on NAFTA. Stegmueller (2013) implements a Monte Carlo experiment to compare frequentist and Bayesian approaches in the determination of the number of countries in a multilevel (hierarchical) model. Additionally, Hare et al. (2014) carries out Bayesian analysis of the Aldrich-McKelvey scaling to analyse American citizen's ideological preferences.

The outline of the paper is as follows. In Section 2 we review the political background that gave rise to the creation of INE. In Section 3, we describe the details of the INE quick count. In Section 4 we present the model used to process the data and show the results obtained in 2006 for the presidential election. Finally, Section 5 concludes with some final remarks.

2. Political background

For more than seventy years (1929–1994) the presidential elections in Mexico were won by the political party currently known as *Partido Revolucionario Institucional* (PRI). Only a few years after the revolution war (1910–1917) this party was created as an instrument for which the groups that emerged triumphant should organize themselves to share the power. For the general population, there was no distinction between government and party. In fact, nowadays that structure, which claimed to be a democratic system, is known as a "unique party regime". Any opposition was essentially annihilated by means of physical violence, all kind of threats, and many forms of bribery.

Only after a particularly well documented repression of students' political movement in 1968, when the army was used against demonstrations with the result of many deaths, in the 70's the government decided to allow the real existence of other political parties. As a consequence, a small number of organizations were able to participate in the electoral processes, although only in a marginal fashion since all aspects were under strict control of the government. For example, the approval of new parties and the regulation for the existing ones was entirely in hands of the Ministry of Interior. Even more, every election was also organized by the same Ministry.

At a very slow pace, Mexico evolved to a system which, nowadays, is close to its counterparts in many democratic countries. Now the elections are organized by INE, an autonomous body with a Board of Directors (*Consejeros*) appointed by the House of Representatives (*Cámara de Diputados*). Besides the role of organizer, INE also acts as a referee among the parties and can impose sanctions if anyone of them breaks an electoral rule. In addition, there is an special Court of Electoral Justice (*Tribunal Electoral*) where the parties can submit any complain regarding other parties, any government officer or even INE itself. It is interesting to know that for the presidential election, INE is, among other duties, in charge of providing public funds to the parties, to supervise the campaigns, to organize the logistics of the election, to hire all the required personnel, to train these staff as well as the citizens who collaborate as officers in the polling stations, to collect and count the votes and to announce the results. However, the final declaration of a winner, if there is one, does not come from INE. Since the parties might find reasons to complain, before or during the process, the winner is only announced by the Court of Electoral Justice after the analysis of every submitted complain.

Remarkably, as a result of the first presidential election organized by INE, in 2000, the office was won by an opposition (right wing) candidate. The same party won six years later, in 2006, when an amazingly closed election took place and a leftist opposition candidate ended in second place. In 2012 the old PRI party came back to the presidency.