Pre-Election Poll Estimations in Mexico: In Search for the Main Sources of Error[1]

Alejandro Moreno[2]
Rosario Aguilar[3]
Vidal Romero[4]

Abstract: In this paper we test different hypotheses that reflect some of the most common sources of estimation error in pre-election polls. We test for questionnaire design effects, sampling effects,

---

[1] This paper reflects the discussions and ideas of several professionals who took part in various meetings with the Federal Elections Institute (IFE) since October 2010, when IFE organized a conference in Cocoyoc, Mexico, to discuss why most pre-election polls were wrong in their estimations for several state-level governor races that year. In that conference, Michael Traugott, a past President of WAPOR, called for a collective effort to design a more formal study of estimation error and its causes. The proposal was backed by IFE councilors and by representatives of the main survey research associations in the country: AMAI (the leading market research association in Mexico), the national representation of WAPOR, and the Consejo de Investigadores. The authors would like to thank the initiative and comments of several participants in those meetings: IFE councilors Leonardo Valdés, Arturo Sánchez, and Benito Nacif, and IFE personnel, Arminda Balbuena Cisneros, Andrea Foncerrada, and Palmira Tapia. We thank our colleagues from AMAI: Alex Garnica, Gregorio de Villa, Gabriela de la Riva, and Edmundo Berumen; from the Consejo de Investigadores: Francisco Abundis, Ricardo de la Peña, and the late Daniel Lund; from the Center for Social and Public Opinion Studies at the Mexican Congress (CESOP): Ángeles Mascott and Gustavo Meixueiro; and the following academic researchers and practitioners: Irma Méndez, Julia Isabel Flores, Federico Estévez, Javier Aparicio, Ulises Beltrán, Alejandro Cruz, Roy Campos, and Jorge Buendía. A previous version of this paper was presented at the 64th Annual Conference of the World Association for Public Opinion Research, WAPOR, Amsterdam, The Netherlands, September 21-23, 2011.

[2] Department of Political Science, Instituto Tecnológico Autónomo de México, ITAM, Mexico City, amoreno@itam.mx

[3] Division of Political Studies, Centro de Investigación y Docencia Económicas, CIDE, Mexico City, rosario.aguilar@cide.edu

[4] Department of Political Science, ITAM, Mexico City, vromero@itam.mx

interviewer effects, spiral of silence effects, and several contextual effects (such as the perception of safety or danger in a polling point in face-to-face polls). We analyze data from a state-level pre-election poll conducted in the State of Mexico on June 2011, two weeks prior to election-day. This poll included an embedded experiment about the placement of the voting question and recorded several contextual variables that allow us to test for different possible sources of estimation error. In addition, this paper offers a brief review of pre-election polling in Mexico during the last two decades, evaluating the polls' performance in both national and state-level elections. This analysis is part (and certainly the first formal step) of a larger effort by polling firms and public opinion researchers, as well as by the Federal Elections Institute, to determine the most common causes of estimation error in Mexican pre-election polls.

Key words: Pre-election polls in Mexico, sources of estimation error, experimental designs, spiral of silence effects, polling methodologies.

Resumen: En este artículo testeamos diferentes hipótesis que reflejan algunos de los errores de estimación más frecuentes en las encuestas preelectorales. Testeamos los efectos del diseño de los cuestionarios, del muestreo, del entrevistador, de la espiral de silencio y de diversos efectos contextuales (tales como la percepción de seguridad o de peligro en los lugares en que se realizan las encuestas cara a cara). Analizamos datos provenientes de una encuesta preelectoral realizada en el estado de México en junio de 2011, dos semanas antes de la elección. Esta encuesta incluye un experimento sobre el lugar de la pregunta sobre el voto en el cuestionario y registra distintas variables contextuales que nos permiten testear posibles fuentes de error en las estimaciones. Por otra parte, este artículo incluye una breve revisión de las encuestas preelectorales en México en las últimas dos décadas, y evalúa la performance de las encuestas tanto a nivel nacional como estatal. Este análisis es parte (y ciertamente un primer paso formal) de un esfuerzo mayor de las empresas encuestadoras y de los investigadores de la opinión pública, así como del Instituto Federal Electoral, por determinar las causas más comunes de errores de estimación en las encuestas preelectorales mexicanas.

Palabras clave: encuestas preelectorales en México, fuentes de error en las estimaciones, diseños experimentales, efectos de la espiral de silencio, metodologías de encuestas.

Introduction

What are the causes of estimation error in Mexican pre-election polls?[5] This question may have several possible answers, but none of them was convincingly provided in 2010, when the Mexican polling community failed to accurately estimate (and foresee the result of) several state-level governor races. Some of these elections were complicated, indeed, as they confronted what was called an "unnatural alliance" of left and right parties, the PRD and PAN, against the old ruling party, PRI. In those states—Oaxaca, Puebla, and Sinaloa—the PRI continued to dominate even after alternation took place in many state governments since 1989 and in the national government in 2000. The majority of public polls predicted a PRI victory in the three states, but the left-right opposition alliances ultimately won. A myriad of explanations were discussed, but none of them presented any evidence to support them. "Spiral of silence effects" was a popular argument this time (called *voto oculto*, or hidden vote), since those states were—according to most political analyst—"authoritarian" environments in comparison to other freer and more competitive states. If true, this may have been the case of Oaxaca and Puebla, but not so much in Sinaloa. Other explanations focused on sampling. Oaxaca and Puebla have substantially rural, poor, and indigenous populations, but sampling issues had not been a problem in those same states in previous elections,

[5] By "estimation error" we mean the difference between the survey estimation and the result of the election.

and it was certainly not the case in other similar states like Chiapas (another relatively poor and rural state with a significant indigenous population), where pre-election poll estimates have been relatively accurate in past years. Others pointed out problems of defective interviewer training and supervision, but many of the polling firms that were wrong in 2010 have also published accurate pre-election poll estimates before and after that year, making it hard to accept that as a final explanation. Some political observers argued that the 2010 poor performance of polls was the beginning of their decadence (see Estévez 2010), unless some methodological adjustments were made. But which adjustments were appropriate? Not a single answer accompanied by systematic evidence was provided. Some suggested, for example, that face-to-face polls (which are still the great majority of election polls in Mexico, where there is a limited coverage of residential telephone lines) should no longer use a secret-ballot method and ask voting preferences directly, but no evidence was provided to support this suggestion—the polling community is not even convinced that this could be the main problem.

The errors of 2010 were an alarming experience for the polling profession in Mexico and a plate full of candy for its critiques. The polls failed to predict the winners and pollsters lacked a clear and solid explanation for this failure. This circumstance fed strong suspicions of data manipulation by some pollsters. The decadence argument kept playing during the months after the election as a requiem in a funeral march, in which the polls' credibility was carried away slowly. Nonetheless, several state-level elections were relatively well-predicted by pre-election polls only a year later, in July 2011. It may now seem clear that 2010 could have been an outlier of generalized errors. But the explanations of why this happened are still missing. As we will argue below, most elections in Mexico (at

the national and state levels) have been accompanied by good and bad polling estimations; a few elections have had almost every poll right, and virtually none, until 2010, had had a majority of polls wrong.[6]

In this paper we aim to contribute to the search for explanations about polling accuracy (or lack of) in Mexico. This topic cannot be fully understood and documented with one paper or with the limited empirical evidence that we have at hand; but we hope that this paper begins a more formal and ambitious search by testing some of the hypotheses that pollsters tend to use when explaining why polls get it right or wrong, and switch to a more substantive debate that is based on empirical evidence rather than on beliefs or guessing Our objective is to provide evidence for a few hypotheses derived from a longer list and build upon the evidence as it accumulates in the road to the 2012 presidential election. We hope that our findings contribute to the development of more experimental design in future elections and certainly in the 2012 presidential election.

The hypotheses that we test in this paper have to do with both methodological concerns and theoretical expectations about voters' responses and behavior. We test, for example, possible effects caused by the placing of the voting question, we comment on the use of a secret-ballot method vs. a direct question, and employ proxies for last-minute changes of preferences and spiral of silence effects. We also analyze how the screening of likely voters works using different approaches, and we assess the impact of interviewer characteristics and supervision on the polling estimates. Our list of hypotheses even incorporates the increasing difficulty that interviewers face in an environment of crime and drug trafficking. For example, interviewers

---

[6] See Romero (forthcoming) for an evaluation of the 2010 polls and testing of some hypotheses.

in the State of Mexico poll recorded the perceived risks and potentially unsafe situations in their polling point. A more formal listing of hypothesis and their justification is discussed below, as well as the primary poll that we used for this analysis. But before we get into the details of our research, it may be useful to summarize the experience of pre-election polls in the country, as a context for further discussion.

Pre-election Polling in Mexico: A Brief Summary

Political scientist Adam Przeworski (1991) once said that a democracy is a system where parties lose elections and also a system where electoral outcomes reflect the institutionalization of uncertainty. The establishment of a dominant-party regime, or a hegemonic party system, as Sartori (1976) called it, made elections a rather certain aspect of Mexican politics for many decades: the PRI would always win and there was no uncertainty about it. Consequently, pre-election polls were virtually inexistent until the 1980s, when the country began to experience increasing political competition and even to witness PRI defeats at the local level, especially in state capital cities and other important urban centers. In the 1988 presidential election, there were there were very few national pre-election polls. At least one of them was not meant for the Mexican public, but for Spanish-speaking television audiences in the United States, while another one found some resistance to its publication in an environment where the media were still under tight government control (see Basáñez 1995). Since then, the country has witnessed a rapid process of political change, driven by several electoral reforms that have re-shaped Mexican politics, by a substantial increase in party competition, and by voters' realignment into several

political options from what once was a single party regime. National, state, and local elections have become, in many instances, very close contests. In a changing political context like this, pre-election polls have established themselves as a normal component of the election campaigns, not only increasing their number in presidential elections but also their presence in state and local races.

Table 1. Final pre-election polls in Mexico's national elections: Number of polls, average errors, and standard deviations

	Number of "final" pre-election polls	*Average error of all polls (Mosteller 3)*	*Standard deviation*
Presidential elections			
1994	9	3.10	1.85
2000	12	2.77	1.07
2006	16	2.38	1.04
Mid-term legislative elections			
1997	6	2.79	1.11
2003	10	2.82	1.10
2009	8	1.64	0.70

Source: Moreno (2009) for the elections prior to 2009, and authors' calculations for 2009, all based on *Reforma*'s record of pre-election polls.

As an illustration of this, in every presidential election since 1994, the number of national-level pre-election polls has increased[7]. According to a count made by Moreno

[7] Presidential elections in Mexico are held every six years and mid-term legislative elections every three years. Elections for state governor in the 31 states and the Federal District also take place every six years and there are elections of this type almost every year (a few exceptions have taken place in the last decade and a half or so that reflect changes in state legislation and scheduling).

(2009a), the number of polls conducted and published between two and three weeks prior to the presidential elections, in most cases by different polling organizations, was nine in 1994, twelve in 2000, and sixteen in 2006. This is a count that considers only what is called the "final pre-election poll".[8] The number also increased from six to ten polls from the 1997 to the 2003 mid-term legislative elections, according to that count, and decreased slightly to eight in 2009—perhaps as a result of the financial crisis that year. Moreno's count also provided some raw estimates for the polls' accuracy. Considering a Mosteller 3 method, the average error for all final polls published decreased from one presidential election to the next: an average of 3.10 in 1994, 2.77 in 2000, and 2.38 in 2006.[9] This can be seen in Table 1.

A somewhat similar trend is observed in the mid-term legislative elections, in which the most recent estimations have been the most accurate: as shown in the lower part of Table 1, the total average error was 2.79 in 1997, 2.82 in 2003, and 1.64 in 2009. Considering that the standard deviations of these average errors may be a reflection of how consensual or different the pre-election estimates of the different polling houses are, we observe that the standard

---

8 Several polling firms conduct a series of polls throughout the campaign. No tracking polls were published in those years, even though they were being conducted for parties or media.

9 We use Mosteller 3 because it offers a standardized way to deal with the Mexican multi-party system, in which at least three parties tend to obtain a significant percent of the vote share, and still be comparable to the state-level cases with two-party local systems. The average approach in Mosteller 3 also allows us to add the standard deviation as a measure of collective poll performance in a relatively understandable way. Mosteller methods 3 (the absolute average error) and 5 (the difference in winner's lead) are common methods for the assessment of poll accuracy, as the works conducted for the United States and Portugal have shown in recent years (see Traugott 2005; and Magalhães 2005).

deviation has also decreased; this means that there has been less disagreement between the different polling estimates. If a trend could be established from this short history of polling at the national level, we could argue that in Mexico the number of pre-election polls in national elections has increased, their overall accuracy has also increased, and the variance in estimation between the different polling organizations has decreased. (More polls, more accuracy, less disagreement).

This information reflects the story of pre-election polling in national elections, but what about polling in state-level elections? At this level the information is less systematic but it is possible to still reach some conclusions. The data at hand give us a good idea of how different the performance of polls has been at the state-level in comparison to the national level, using the collection of data gathered by the Department of Public Opinion Research at *Reforma* newspaper since 1999. This collection consists of the final pre-election estimates for both *Reforma* and all other polling organizations that *Reforma* personnel was able to detect during the course of a campaign.[10]

There are at least three limitations and one warning in regards to this data collection. The first limitation is that competitive elections at the state level go back to 1989, when the PRI lost the first state-level race since 1929. So this collection of data begins ten years after the first alternation in a governor's office—still, the number of polls then was much more reduced. A second limitation is that the data only reflect those elections where *Reforma* conducted state-level pre-election polls and published a

[10] We thank Yuritzi Mendizábal and Rodrigo León for preparing this information, and to *Reforma* for making it available for further research.

"final poll"[11]. Despite of this, the public opinion research unit at *Reforma* is one of the "polling houses" that has one of the largest coverage in the country when it comes to conducting pre-election polls, covering 78% percent of the state-level elections that took place from 1999 to 2011.[12] This is undoubtedly one of the most comprehensive collections of pre-election poll data that is available at the time of writing this paper.

A third limitation is that the gathering of information was conducted from Mexico City, capturing the numbers and estimations of different polling houses that are nationally known, and some local polls that were "visible". The potential problem with this is that some local polling organizations and the publication of their work may have escaped this data collection. The warning is that the data collection by *Reforma* includes poll estimations by well-known as well as by unknown polling houses, polls reported in journalistic stories, polls reported in political columns, and polls published as political advertising. There was no discrimination of the entries; but each poll was counted only once. Having said this, let us see what the state-level

---

11 Our emphasis on "final polls" is based on the fact that they are the ones that yield the most accurate projections for the election. An interesting analysis of pre-election poll accuracy in Mexico considering polls published during the campaign and even before candidate nominations can be found in Romero and Varela (2011); they show that the level of accuracy in all polls improves as election-day approaches, as it should be expected.

12 A total of 68 gubernatorial elections were held in the 1999-2010 period, including two extraordinary elections in Tabasco and Colima states. The number of elections in this count also considers two elections for Mexico City's head of government (Jefe de Gobierno), and two elections in Mexico City for the Federal District Legislative Assembly, ALDF, in which *Reforma* published final pre-election polls.

history of polls looks like according to the *Reforma* count and registry of poll results.[13]

Table 2 shows information for 53 state-level elections held from 1999 to 2011 in which *Reforma* conducted and published a final pre-election poll. In those same elections, the newspaper's polling unit recorded 217 final pre-election polls conducted by various polling organizations, including its own. These final pre-election polls had a total average error of 3.47 for the entire period (an average of averages) and a standard deviation of 2.33. The table also shows two periods that correspond to the electoral cycle, that is, when elections in the same states take place again: from 1999 to 2004 (six years), and from 2005 to 2010 (six years). This information indicates that the most recent period actually had more polls on average per election, a lower average collective error, and a lower standard deviation (again: more polls, more accuracy, less disagreement). The number of elections considered in the most recent period, however, is lower than in the first (they should be about the same, but this reflects a reduced coverage by *Reforma* in its pre-election polling estimations in recent years). Because of this, we cannot argue with certainty that the trend is towards more polls, more accuracy, and less disagreement. But we can say that the collective average error is very similar in both periods.

---

13 A much more exhaustive count and registration of poll estimates requires the systematization of information reported to the election authorities, both at the federal and state-levels. In Mexico, polling organizations that publish their results are required by election laws (federal and local) to report their methodology, their results and other aspects of their polls to the corresponding election authority. An undetermined number of reports for many years are available at IFE (for national and some local polls) and may be available in every state election institute, which gives a good idea of the titanic task that having this information in a single data set involves.

Table 2. Final pre-election polls in state-governor elections in Mexico: Number of elections, number of polls, average errors, and standard deviations

	Number of elections	*Number of polls*	*Average errors (Mosteller 3)*	*Standard deviations*
Years				
1999-2011	53	217	3.47	2.33
Periods				
1999-2004	30	97	3.51	2.37
2005-2010	21	100	3.33	2.27
2005-2009	17	73	2.58	1.65
2010	4	27	4.81	2.57
2011	2	20	3.99	2.52

Source: Election-specific data provided by *Reforma*'s Department of Public Opinion Polling and authors' calculations. It includes two elections for the Federal District's Legislative Assembly.

Table 2 also shows 2010 separately so we can observe the magnitude of the polling error in that particularly bad year, which increases to 4.81.[14] In 2011, the start of a new election cycle and with only 3 state elections so far (*Reforma* conducted and published final polls in only two of them), the average error of all polls published went down again to

[14] Surprisingly, when we look at the year-by-year average errors for all elections and all published polls, 2010 is not the worst year. In 2000, several polls conducted by unaccounted (and perhaps inexistent) polling organizations, such as Technomanagment, yielded a total average error of 5.10. The average error for that particular "polling house" in the state of Morelos was 12.45! However, the 2000 average was also increased by a bad performance in a single state by a more well-known and reputable polling firm. The difference between 2000 and 2010 is that in the former there were a few very bad polls and several very good ones, whereas in 2010 the bad estimates were more generalized, including the polls from various reputable firms.

an under-4 level. This shows that 2010 could have been, effectively, an outlier of bad polling performance. Regardless of that, the average error of pre-election polls published is greater than the sampling error that polling organizations usually report in their publications (which tend to vary from +/- 1.8 to 3.5 percent). Thus, the need to explain inaccuracy and to have a better understanding of the sources of error that are most common in Mexico remains.

Testing for various sources of error in a state-level pre-election poll

The remaining of this article is devoted to the analysis of how different sources of error in a pre-election poll may affect the polls' accuracy. Since our evidence comes mainly from one pre-election poll and a single election, we will be unable to generalize the results that we obtain, but, again, this is one of the most systematic attempts to understand the source of error in Mexico and what we hope to be a first step of a more systematic study of poll estimation error in the country. The discussion is organized in the following parts: a description of the data; a listing of hypotheses; a presentation and discussion of results; and a concluding section in which we point out some of the topics that could be covered in future research.

Description of data

Unless otherwise indicated, our analysis is based on an actual state-level pre-election poll conducted by *Reforma* newspaper, in which we had the opportunity not only to include an experimental design on the placement of the voting question, but also to record several items that are

relevant to the testing of various hypotheses on estimation error. Such items are derived both from the respondents' answers and from interviewers' and supervisors' observations. The poll was conducted in the State of Mexico on June 18-19, 2011, two weeks before the election for state governor, scheduled for July 3. As in many other states, the State of Mexico's election legislation forbids the publication of public opinion poll results within eight days prior to the election and until after the polls close, which usually forces most polling houses to conduct their final pre-election poll for publication up to two weeks before election-day. The poll was conducted face-to-face in 80 polling points probabilistically selected from the list of precincts or electoral sections. Blocks and households were also probabilistically selected and then, in the last stage of selection, respondents were selected using quotas of sex and age.[15] Fifteen interviews were conducted in each polling point. The refusal rate for the poll was 24%.

The poll results were published by *Reforma* on June 22, with seven days of campaigns still ahead and eleven days before the election. Actually, a televised debate between the three candidates for governor was held on the evening of June 22. Some observers argue that last minute campaign effects not captured by the final pre-election polls conducted as early as the election laws allow may in fact explain a great deal of the polls' inaccuracy (see Estévez 2010). However logical this assertion is, no empirical evidence has been gathered to support it in Mexico. The closest attempt is perhaps the journalistic reports from exit polls about the percent of voters who said they made up their

[15] A source of error that we are not able to test using this poll is the effect of quotas, since all respondents were selected using that criteria. In the sampling issues discussed below, we acknowledge the need to test for selection methods in the sample.

minds in the last few days before the election and even on election-day. In 2006, for example, a national exit poll showed that the largest proportion of late deciders chose the PAN presidential candidate, who ultimately won the election by only half percentage point (see Moreno 2009b: 245). For 2010, Romero (forthcoming) finds no evidence of systematic biases on polls due to last minute events. In other countries evidence shows that a proportion of the electorate makes up their minds within two weeks of election-day (for the U.S. see Zaller 2004; ANES 2007. For the German case see Schmitt-Beck & Faas 2006. For the French case see Reuters 2007). Based on exit poll data Nir and Druckman (2008) conclude that those who decide their vote towards the end of the campaign are ambivalent voters who received mixed-information messages from the media in a highly competitive race. There is ample room for conducting more research in this area, but what we know now is that whether people will make a late voting decision depends on the type of election and media coverage. In fact, the prohibition of publication within a certain number of days from the election has been in effect for many years in Mexico, and all polls considered in the review of poll estimates in the previous section were conducted within this regulation. In many cases, the poll average error has been close to zero or well under one percentage point, so the claim that last minute changes may affect accuracy cannot be generalized. However, it certainly is a credible possibility in some elections, like the 2006 presidential race.

Table 3. Comparing the State of Mexico election results with the pre-election poll estimates under various treatments

	PRI	PRD	PAN	Average error (Mosteller 3)	Difference 1st – 2nd
	%	%	%		
Official election results (99%)	65	23	12	--	42
Poll results:					
With no treatment	63	23	15	1.7	40
Screened by likely voters (individual motivation)	66	21	13	1.3	45
Screened by likely voters (social or family motivation)	66	20	14	2.0	46
By voting question placement (near the beginning)	64	23	13	0.7	41
By voting question placement (near the end)	62	22	16	2.7	40
Combining likely voters screening (individual) and question placement (beginning)	69	20	11	2.7	49
Combining likely voters screening (individual) and question placement (end)	64	21	16	2.3	43
Filtered by likely voters and weighted by 2005 vote	59	26	15	4.0	33

Sources: State of Mexico Election Institute (IEEM) for the official election outcome; and *Reforma*'s final pre-election poll in the State of Mexico (see methodological details in the text).

Table 3 shows a comparison of the election official results with the poll estimates. The latter are shown in different versions, as a way to start looking, if not quite at sources of error, at least to the possible effects of different treatments employed. The first row shows the election results, followed by the raw poll results without any treatment. By treatment we mean weighting techniques or screening by likely voters. Voting is compulsory in Mexico but there is no sanction to those that do not turnout to the polls, and consequently turnout rates tend to be lower in comparison to other Latin American countries. Unless they take place at the same time of a national election, many state-level races actually have relatively lower turnout rates (between 40 and 60%), and the State of Mexico is no exception. On July 3, 2001, the state registered a turnout of 56% in the election for governor. Under these scenarios of low to medium levels of turnout, a screening of likely voters may help increase the accuracy of poll estimations.

As shown, the raw poll results had an average error of 1.7, with errors of two and three points for the PRI and PAN, respectively (the poll's reported maximum sampling error was +/-2.8%). The table shows two different treatments in terms of likely voter screening. The first is based on individual motivation variables (basically the respondent's interest in the campaigns and his/her subjective probability of going to the polls, questions asked immediately before and after asking if the respondent knew the date of the election). As shown, screening for likely voters by the use of these individual motivation variables resulted in a slightly better estimate (a 1.3 average error and a maximum error of two points for the PRD) than the raw results. Nonetheless, the difference between first and second place (shown in the rightmost column) was actually widened by one absolute point, so the results were more accurate if

we use Mosteller 3, but not if we use another indicator of accuracy (like Mosteller 5).

A second mode of likely voter screening that this poll offered reflects social or family motivation. This was based on the respondents' perceived political environment in his/her household, that is, how much she/he reported that the family members talk about the candidates and their campaigns, and what is the respondent's perceived probability that his/her family will turnout to vote on election day. Unlike the individual motivation variables, these family environment variables are based on the logic that voting is a social act and that Mexicans may be to some extent influenced by the environment they perceive at home. Other aspects such as the mobilization by parties or by other secondary groups were not considered in this poll.

As shown, the estimation based on screening of likely voters by social/family motivation was not as accurate as the estimation that used screening of likely voters based on individual motivation. The average error in the social or family model was 2.0, and the difference between first and second place widened to four points (as opposed to three points in the individual motivation mode). In any case, the most noticeable error in both modes of likely voter screening had to do with underestimation of support for the leftist PRD, but estimates of both PRI and PAN improved in comparison with the raw poll results.

Table 3 also shows the poll results broken down in two groups derived from an experimental design. In half of the interviews the voting question was asked near the beginning of the questionnaire (it was the sixth question out of 25), and in the other half it was asked later in the questionnaire (it was question number 24).[16] This treat-

[16] The voting question near the beginning was preceded by the perceived main problem in the municipality where the respondent lives, the

ment was applied alternating the type of questionnaire, one to one respondent and the other to the next. It was not applied by polling point, which would also have been useful for the type of analysis that we develop below. This experiment was conducted as a way to assess whether the placement of the voting question (before or after the respondent's thinking about campaign issues) contributes to higher or lower levels of accuracy of the poll estimates.

As shown in the table, the placement of the voting question near the beginning (and prior to any further political reasoning) yielded a better poll estimate in this case, with an average error of 0.7 (the lowest of all average errors shown in the table), as compared to the average error of 2.7 derived from the late question placement. Without any possibility to generalize, this results suggest that an early and "clean" question about voting preference may be more helpful than one asked later and after other items that involve further political or economic reasoning.[17] Table 3 also shows the combinations of likely voter screening (by individual

---

individual motivation likely voter items described earlier, and a question that asked if the respondent had already decided his/her vote or whether she/he was still undecided at the time of the interview. The voting question placed later in the questionnaire was preceded by the respondent's opinion about the candidates, exposure to campaign events, to the televised debates and assessments of the winner, the president and governor approval ratings, self-reported crime victimization and loss of job in the household, and confidence in the election authorities.

17 The State of Mexico poll conducted by *Reforma* did not include an experiment about asking voting preference with a secret-ballot method vs. asking it verbally and directly. Nonetheless, a similar pre-election poll conducted by *Reforma/Mural* for the Jalisco state governor race, in 2006, included a somewhat similar exercise, yet not quite experimental. The voting question was asked at the beginning using a secret-ballot method and again at the end asking verbally and directly. The results from the question placed at the beginning with a secret-ballot method yielded better estimates of the election result than the question placed at the end which was asked directly. The problem with this poll is that both modes were asked to the same respondents, and there is no way

motivation only, since it proved to yield more accurate estimates in this poll) and the voting question placement (near the beginning and near the end). The combination of these two treatments increased the average error when screening for likely voters and asking voting preference near the beginning, in comparison to the estimates for each of these treatments separately. However, the combination improved the estimates slightly when the voting question was asked later, which means that the overall influence of likely voter screening may be beneficial to pre-election polls, regardless of whether they ask voting preference at the beginning or at the end.

Finally, Table 3 also presents in its last row an estimate treatment that combines likely voters and weighting for the vote in the previous election. Weighting for prior vote is a common practice of pre-election polls in Spain, for example. It is based on the assumption that voting preferences do not change dramatically from one election to the next, an assumption that may not apply to electorally volatile emerging democracies. In Mexico, state-level elections conducted by *Reforma/El Norte* in states where the PRI had a substantial lead, such as Coahuila and Durango, with support well over 60%, the poll results actually overestimated an already high level of support for the PRI. In the State of Mexico this could have been the case, thereby inviting to consider the previous vote as an anchor. Following the experiences of Coahuila and Durango, *Reforma*'s final publication included both the raw results and the results derived from a projection that considered likely voter screening and weighted the results by prior vote (the most recent governor election of 2005) in the State of Mexico. This projection estimated the support

---

to determine what the main cause of the inaccuracy was: the late placement, the direct question, or something else.

for the PRI at 59%.[18] With the latter mode, the PRI was underestimated by six percentage points, and the opposition parties were overestimated by three percentage points each, for an average error of four points and a significantly lower difference between first and second place: 33 points, as compared to the final 42 points. (This illustrates that the influence of prior experiences in a current projection may also be among the sources of estimation error, but they are part of the data treatment and not the poll error in itself).

Table 4. Pre-election poll estimates for the 2011 State of Mexico election published by four newspapers (All face-to-face polls except the GCE telephone tracking poll)

	PRI	PRD	PAN	Average error	Difference 1st – 2nd
	%	%	%		
Official election results (99%)	65	23	12	--	42
GCE/*Milenio* (telephone tracking poll)	62	23	15	2.0	39
Reforma (raw results)	62	23	15	2.0	39
Reforma (projection)	59	26	15	4.0	33
El Universal	59	27	14	4.0	32
GCE/*Milenio* (face-to face)	61	20	18	4.3	41
BGC Beltrán / *Excelsior*	58	27	15	4.7	31

Note: Results are rounded up.

[18] In 2005, the PRI candidate won the election for governor with 49%, whereas the PAN and PRD candidates obtained 26 and 25%, respectively.

All other polls sponsored by newspapers also underestimated—to a greater or lesser extent—the PRI vote in the State of Mexico election, as shown in Table 4.[19] The total average error for five pre-election poll estimates ranged from 2.0 to 4.7. Interestingly, a telephone tracking poll conducted by Gabinete de Comunicación Estratégica (GCE) for *Milenio* newspaper yielded comparatively accurate results—and much more accurate that the same firm's face-to face poll. This raises the question about the accuracy and appropriateness of telephone vs. face-to-face polls. The State of Mexico experience cannot be generalized but in this case there was a difference in the polls results from the same organization. Undoubtedly, discussions about sources of error will have to add the mode of interview in the future (telephone vs. face-to-face). Among the reasons why telephone interviews are seldom used in Mexico for pre-election polls is the fact that residential telephone lines have a limited coverage and the fact that telephone polls have a bias towards higher socio-economic levels, not to mention urban settings, as opposed to rural ones. The telephone interviews fortuitous performance in the State of Mexico in 2011 may be explained by the fact that voters in several middle class districts that usually favor the PAN opted this time for the PRI, but this is just a speculation in our part.

Let us now move on to the analysis of the poll's sources of error. In the following two sections we list a series of hypotheses and then proceed to an empirical analysis with variables that attempt to represent each of the hypotheses proposed.

---

[19] Two newspapers, *Reforma* and *El Universal*, used their in-house polling units for the polls, whereas *Milenio* and *Excelsior* hired or established a collaboration scheme with polling firms GCE and BGC Beltrán, respectively.

Hypotheses (and indicators)

There are several hypotheses that are usually pointed out as common sources of poll estimates error by Mexican pollsters and by observers of the polling profession in the country. In the remaining of the article we attempt to test some of those hypotheses using the poll conducted by *Reforma* in the 2011 State of Mexico election.

The hypotheses that we attempt to test are the following (preceded by an analytical category in capital letters as a way to classify the sources of error):

SAMPLE DESIGN:

As an essential part of polling, sampling is a natural source of error. There are various hypotheses that can be tested under this category, but the data at hand (which recorded whether the interview was an original selection or a substitution) allow us to test the following:

1. *Substitution of the original sample respondents, for whatever reason related to sample non-response, increases the estimation error.*

In addition, we also test for the difference between urban and rural samples. We have no a priori expectation about this, although some pollsters argue that interviews in rural areas tend to yield less accurate estimates.

QUESTIONNAIRE DESIGN:

How the questionnaire is designed and how the questions are asked is also a natural source of error, in polls. In this analysis we test whether placing the voting question before and after the items that activate political and economic reasoning affect the poll estimates.

1. Voting question placement affects the accuracy of poll estimates.

1.1. Asking for voting preferences at the beginning of the interview increases poll accuracy because it measures a more spontaneous and "cleaner" response without the possible influence of other items in the questionnaire.

1.2. Asking the voting question later in the questionnaire allows the respondent to take several factors into account during the interview before revealing his/her preference. The possible bias from this reasoning may be positive (1.2.1) or negative (1.2.2), thereby increasing or decreasing the accuracy of poll estimates.

SPIRAL OF SILENCE EFFECTS:

One of the favorite, most simple and most common hypothesis that Mexican pollsters use refers to situations where respondents don't reveal their real preferences for some reason ("voters lie", some pollsters say). Spiral of silence effects is a more formal way to represent this on the basis of a theory of survey response (Noelle-Newmann 1974). Nonetheless, it is not an easy task to test for these effects. In this analysis we propose a way to do it according to the following hypothesis.

2. Spiral of silence effects take place especially when the respondent perceives him or herself to be among the minority view. This perception may lead him/her to give socially desirable responses during the interview, including the vote preference.

2.1. Perceiving oneself among the minority view increases the probability to give a socially desirable response and, therefore, to a higher inaccuracy of the poll estimates.

2.2. Spiral of silence effects may be more noticeable when the respondent perceives that he/she is among the minority view in his/her closer community (where the pressure of social norms is higher) than when he/she perceives him/herself to be among the minority in

a broader and more abstract community (say, the state as a whole).
For the latter two hypotheses, the State of Mexico poll included two items right after the voting question that inquired whether the respondent believed his preference to be among the minority or the majority view in his municipality and in the state as a whole.
2.3. Spiral of silence effects are also present when the respondent perceives a potential pressure from the interviewer.

The *Reforma* poll in the State of Mexico also included a record of possible pressure felt by the respondent: Was there any situation during the interview in which the respondent felt pressure or did not feel free to express his/her opinions? Was there a moment when the respondent felt distrust towards the interviewer? Did the respondent think that the interviewer was working for a political party? An indicator of pressure was constructed from these indicators, which were coded by the interviewer at the end of each interview.

CONTEXTUAL EFFECTS:

The respondent in an interview (and also the interviewer) may feel pressured or threatened by contextual effects, which refer to the general environment where the interview takes place. The increasing violence and crime, for example, have made face-to-face polls increasingly difficult in Mexico.[20] An unsafe environment not only may affect the quality of an interview (and the responses) but

[20] WAPOR's press release in early August 2011 about the disappearance of interviewers working for the Consulta-Mitofsky and Parametría polling firms in the State of Michoacan, Mexico, is a good illustration of this problem.

also the work done by the interviewers, so it is a potential source of error. We hypothesize that:

3. An unsafe or threatening environment in a polling point decreases the accuracy of poll estimates.

INTERVIEW(ER) EFFECTS:

Among the many sources of error identified by Herbert Weisberg (2008) in his monograph of total survey error, interviewer effects are quite important. We also include the interview itself here as a way to have a broader set of possible sources of error. In this paper we test the direct interviewer effects (in this case represented by the sex and by the age of the interviewer, as well as by his/her interviewing experience); the effects of direct supervision; and the effects of the length of interview. We hypothesize that:

4. The sex and age of interviewers may affect the poll estimates by increasing or decreasing the respondent's confidence in the interview. We believe that female interviewers may have a more positive influence than male interviewers, especially in an environment where crime has increased, as it is the case in Mexico (as female interviewers may seem less threatening than male ones).

5. The interviewers experience may also affect the poll estimates. We hypothesize that more experienced interviewers are more likely to increase poll accuracy than less experienced ones.

6. Direct supervision of the interview may contribute to increasing the accuracy of poll estimates. We hypothesize that interviewers under close watch tend to do a better job during their interviews. Of course, it is also feasible that direct supervision may increase pressure and affect negatively the poll estimates.

7. The length of interview affects the quality of response. We hypothesize that interviews that take longer than average (for whatever reason), tend to wear out the respondent, contributing to a decreasing quality of his/her responses. This hypothesis has different implications depending on the placement of the voting question. If our expectation is true, longer interviews should have a greater negative impact when the voting question is asked near the end of the questionnaire.

"LAST MINUTE" CHANGE EFFECTS:

Given the relatively long periods that span between the fieldwork of a final pre-election poll that can be published and the election day, forced by the prohibition to publish poll results, it is commonly argued that when polls are not very accurate it is because last minute changes of preference take place. We believe that even though in some cases this may be true, it does not always happen in Mexican elections. Much of the polling history in the country has been characterized by this legal restriction, and yet many pre-election polls achieve a high degree of accuracy, in some case not just considering the final poll, but starting with other earlier polls conducted throughout the campaign.[21] This hypothesis is difficult to test with the same pre-election polls that are constrained by the publication prohibition, but we can employ a proxy for this phenomenon: the undecided voters. They are the most likely to change (or form) a preference in the last few days prior to the election (and they are also the most likely to abstain). In the 2006 presidential election, for example, a small yet substantial proportion of voters

[21] The State of Mexico governor race of 2011 is a good example of how voter preferences changes very little during the campaign, as both polls and tracking polls showed almost null variation.

said in a national exit poll that they had decided to vote for Felipe Calderón in the few days prior to the election, and certainly after all public polls were already published (see Moreno 2009b).

We hypothesize that:

8. "Last minute" changes of preference may take place and affect the poll estimates. (In this case, our proxy for "last minute" changes is represented by respondents who declared themselves as undecided).

Other hypotheses that are important but that we cannot test in this paper with the data at hand are the following, hoping that they can be tested in future studies:

9. The use of a *secret-ballot method* vs. asking voting preferences directly influence the interviewer's response.

9.1. A secret-ballot method helps the respondent feel more confident about the anonymity and confidentiality of his/her response, thereby increasing the accuracy of the poll estimates.
9.2. A direct question seems as a less informal and official act during the interview and helps the respondent express their preference under less pressure.

10. The *screening of likely voters* makes a difference in the poll estimates. There are two opposing views about likely voters in Mexico:

10.1. Screening for likely voters is crucial in Mexico because voting is compulsory but without a sanction, turnout rates are comparatively low and, in addition, turnout rates have been decreasing over time. Screening for likely voters improves poll accuracy in Mexico.
10.2. Screening for likely voters is useless in Mexico because people's responses about the likelihood of voting are not reliable and they do not take into account the mobilization of parties on election day. Screening for likely voters does not affect (or may even decrease) poll accuracy in Mexico.

Although different items for the screening for likely voters were included in the State of Mexico poll, as discussed earlier, for their proper testing an experimental design is needed, where half of the polling points use likely voter items and half do not. (We were not able to do this experimental analysis but it would be possible to test this hypothesis by randomly dividing the polling points into two groups. For the control group the analysis can report the raw results of electoral preferences, while for the treated group the analysis could apply the screening questions to report these preferences).

Results

The results of our analysis are shown in Table 5. This table shows the results of OLS regression in which the dependent variable is the total error per polling point, that is, the total difference in absolute terms between the election official result in the precinct (or electoral section) and the result obtained by the poll in that polling point.[22] We will comment first on each independent variable as it appears in the table, and then we will make more general comments about this analysis. Because the dependent variable is measured as the total error in absolute value, larger values represent a bigger error, and zero means no error. Consequently, positive signs of the coefficients represent a positive contribution to bigger error, and negative signs a contribution to accuracy (or smaller error). The results are shown in four columns, one for the total average error

[22] A single polling point certainly cannot not be a representative sample of a larger precinct, but it should reflect a general trend. In this analysis we assume that the polling point error can be a matter of sampling, and yet assess the effects of other variables in the total error.

for the three political parties, and the other three for each of the parties separately.

Let us discuss the variables that represent sampling effects. First, the substitution of original sample respondents did not have a significant effect on the poll's total error. In this poll, about 24 percent of the original sample (which was unable to contact or who refused the interview) was substituted. As the results show, this substitution did not contribute to increasing the poll estimates' error. Secondly, and against prior belief, the rural sample actually increased accuracy of the estimation of support for PRI and PRD. This had an incidence in the overall accuracy shown in the first column. This is perhaps explained by the fact that the PRI vote was high in all settings but proportionately higher in the rural settings, whereas the PRD vote was proportionally higher in urban areas, particularly in municipalities that are part of the Mexico City metropolitan belt. The PAN did not show any substantial differences in support by urban-rural setting, which is interesting for those who know the electoral history of the state, where PAN candidates have traditionally drawn more support in the metropolitan belt as well.

Table 5. Testing various hypotheses for poll estimation error: OLS regression

	Total absolute error		*Absolute error for PRI*		*Absolute error for PAN*		*Absolute error for PRD*	
	t	Sig.	t	Sig.	t	Sig.	t	Sig.
Sampling effects								
Substitute sample	0.93		0.49		0.41		1.48	
Rural sample	-2.55	*	-2.87	**	1.28		-4.57	***
Questionnaire design								
Voting question at the beginning	-0.37		-0.04		-0.82		-0.08	
Spiral of silence effects								
Minority view in municipality	0.19		-0.40		2.35	*	-1.60	
Minority view in the state	-0.19		0.40		-1.53		0.65	
Pressure interview	3.11	**	3.40	**	1.67		1.92	*
Contextual effects								
Unsafe polling point	-1.38		-0.25		-3.79	***	0.81	
Interview(er) effects								
Female interviewer	-4.60	***	-5.27	***	-5.68	***	1.31	
Age of interviewer	-1.08		-4.09	***	2.99	**	-0.40	
Experience of interviewer	-0.83		-0.34		0.75		-2.79	**
Supervised interview	-3.14	**	-2.99	**	0.61		-5.32	***
Length of interview	0.87		0.21		0.74		1.35	
Last minute change effects								
Undecided	-1.44		-0.99		-1.80		-0.59	
(Constant)	13.22	***	12.80	***	7.73	***	9.95	***
ADJ R-SQ	.04		.06		.05		.05	

Significance levels: * p<.05; ** p<.01; *** p<.001.

The fact that we did not observe significant sampling effects in this analysis does not mean that they are not present. More research that takes different sampling issues into account should expand our evaluation of the sampling effects on poll accuracy.

The placement of the voting question at the beginning seemed to have contributed to more accuracy than the voting question placed later in the questionnaire, but this effect was not statistically significant when controlled by other factors. In this sense, questionnaire design effects on accuracy or inaccuracy were not observed at the aggregate level; nevertheless there are significant effects if we dissagregate the sample by demographics

We find significant differences among specific segments of the population. Table 6 shows the difference on voting preferences between the group interviewed with the voting question at the beginning of the questionnaire and the group with the voting question later in the questionnaire. Men are positively influenced towards the PAN candidate when the voting question is placed at the middle-end of the questionnaire. Women significantly decrease their item non-response rate when the voting question is located at the beginning of the questionnaire.

Table 6. Differences on preferences due to voting question placement (beginning minus middle placement) by sex

	Men	Women
PAN Candidate	-6.1**	1.5
PRI/PVEM/NA Candidate	4.2	2.7
PRD/PT/Conv. Candidate	2.6	2.2
NA	-1.3	-5.4**

We also verified for effects by education. One would expect that questionnaire effects would be bigger on the

less educated, since this segment of the electorate should have less information and is more reliable on other people's opinions to decide. Table 7 shows the differences between both groups in the sample. The largest effects—although not strongly significant—are on the PRI candidate, if asked about their electoral preference at the beginning of the questionnaire, those with no formal education and those with college education were much more likely to vote for the PRI, which is a peculiar result. The effect on uneducated citizens is twice of what we observe for undergraduates. The opposite effect that we find is on the PAN candidate among citizens with elementary school.

Table 7. Differences on preferences due to voting question placement (beginning minus middle placement) by formal education

	PAN	PRI	PRD	NA	Abs. Avg. effect by education
None	-2.1	**20.9***	-2.1	-6.6	7.9
Elementary	**-6.0****	6.4	-1.9	-3.2	4.4
Secondary	0.4	-6.6	4.4	-1.0	3.1
High School/ Technical	-1.9	6.0	4.0	**-4.9***	4.2
College	-4.9	**10.4***	5.1	0.9	5.3
Graduate	12.5	1.1	-20.5	-9.1	10.8
Abs. Avg. effect by candidate	4.6	8.6	6.3	4.3	

Spiral of silence effects, represented by the respondents' perception of being among the minority view, were generally insignificant with the exception of the PAN. The respondents' perception of being a minority in their own municipality actually increased the total error for the PAN estimates. In contrast, perception of being a minority in the overall state did not have any significant effect on the

poll's total error. We looked more closely into whether spiral of silence effects had a greater effect on older people and women as the theory predicts (Noelle-Neuman 1974; Scheufele 2008). For this purpose we ran the same models with multiplicative terms for age, gender, and whether the respondents felt they were among the minority position in their state and municipality.[23] As graphs 1 through 3 show the interactions effects were only significant in three cases: for the PAN when respondents felt they were among the minority view within their municipality, and for the PRD in the case respondents felt among the minority both at the state and municipal levels.[24]

In the case of the PRD the marginal effect of respondents' perception of being in the minority in the state is negative and statistically significant[25] for women between 18 and 31 years old and after 77 years of age when compared to men. In the case of younger women, as expected, their marginal effect on the error term is significantly higher than that for men. Older women behave contrary to the gender expectation, as the marginal effect of feeling among the minority view of the state is less than that of men. Younger male respondents (18-37 years of age) who perceive their preference to be within the minority view in their municipality tend to be significantly less prone to hide their true PRD preference than women from their cohort. In this case, in contrast to the two others, the age expectation (younger people will not care about being among the minority when expressing their opinion) is met, as the marginal effect curves have a positive (instead of a negative) slope.

---

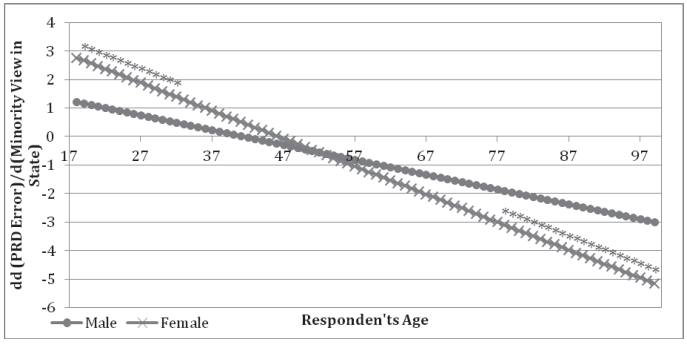
23 The models are in the appendix.

24 In order to find whether interaction effects are significant is necessary to include the main effect of their components. A clear way fo doing it is by presenting the marginal effects of both the interaction and main variables (Brambor et al. 2006; Kam and Franzese 2007).

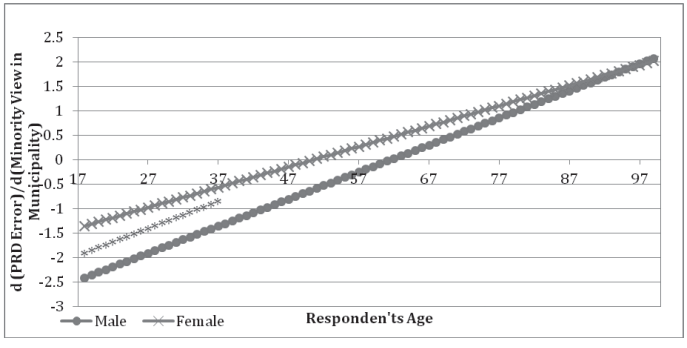
25 Significance levels are set to at a p-value≤0.05

Finally, in the case of the PAN, we can see that is younger men (18-40 years old) who tend to significantly hide their true opinion if they think they are in the minority in their municipality when compared to women of the same age.

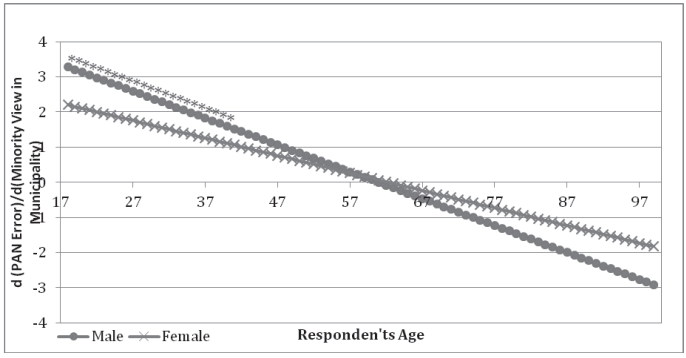
Graph 1
Marginal Effect of Respondent's Perception of Minority Status within the State, by age and gender, on the PRD Error



Graph 2
Marginal Effect of Respondent's Perception of Minority Status within the Municipality, by age and gender, on the PRD Error



Graph 3
Marginal Effect of Respondent's Perception of Minority Status within the Municipality, by age and gender, on the PAN Error



This provides some evidence to our hypothesis that local-level community societal pressure matters more than in a broader and more abstract community (state or country). The only case where respondents' perception of the status of their opinion within the state mattered was for supporters of the party which came last in the elections. We also found that spiral of silence effects differed depending on the age and sex of the respondents. The theory predicts that younger people will feel freer to express their views regardless of whether they felt were with the majority or minority. We found this was just the case for the PRD error when respondents thought they were in the minority at the community level. In terms of gender, depending on the respondents' age, sometimes women would express more their true preference than men. The fact that the parties with the lowest support in the election are the ones for which spiral of silence effects took place is telling about this source of error among actual minorities. According to the poll results, two-thirds of PAN voters believed that the PRI candidate would win the election. Among PRD voters this perception was slightly lower.

Interestingly, the variable that represents pressures during the interview was statistically significant in all cases but the PAN, suggesting that this type of pressures may not quite reflect a spiral of silence effect but a direct interview(er) bias. In the cases of PRI and PRD (although more noticeable in the former), feeling some type of pressure by the interview(er) contributed to increase the poll's error. The more pressure the respondent felt, the more likely he or she said s/he would vote for the PRI. The combination of results from these Spiral of Silence variables suggests that social norms and social pressure may actually have significant effects on accuracy of poll estimates in the country, and more research in this direction should be conducted.

The influence of an unsafe environment in the poll accuracy was generally insignificant, with the exception of the PAN support. As the results show, conducting interviews in polling points perceived as unsafe or potentially unsafe (about 20% of all polling points) increased the accuracy of the estimates for PAN support. This finding goes against our theoretical expectation, that an unsafe environment would increase the polls' inaccuracy. We do not find either higher non-response rates in unsafe polling points or significant differences on estimates for other relevant variables. While it is true that polling is becoming a high risk job for interviewers, we find no evidence that this circumstance is affecting the survey estimates. The subsample of citizens that, despite the unsafe context, chose to answer a poll tends to distribute just like the rest of the sample.

The analysis shown in Table 7 indicates that interview(er) effects did take place in different forms. Let us discuss one by one in the order in which the independent variables appear in the table. First, the sex of the interviewer matters. (It represents the stronger predictor of accuracy in the estimation of PRI and PAN support). Female interviewers contributed significantly to the poll accuracy, as shown by

the negative sign for the coefficients for support for PRI and PAN, but the effect was the opposite and insignificant in the case of the PRD.

The age of the interviewer also shows a statistically significant contribution to accuracy but with mixed results. The sign is negative in the case of the PRI, which means that older interviewers had better results in the estimates for that party; and it was positive in the case of the PAN, with younger interviewers getting better estimates for that party. Our interpretation of this is that identification between the respondent and the interviewer may contribute to better estimates; for example, PAN voters are usually younger than PRI voters. However, the State of Mexico election breaks with this association, since the PRI drew more support among younger voters than it is usually observed.

The interviewer's experience only shows significant effects in the case of the PRD support. A statistically significant coefficient with a negative sign in this case suggests that more experienced interviewers (those who have participated in more polls and have a longer experience doing fieldwork) actually got better estimates for that party. This is an interesting finding, considering that many preelection polls in the past have underestimated support for the leftist party in some instances.

The direct supervision of interviewers also contributes significantly to the polls accuracy, especially in the cases of PRI and PRD. (It results in the largest t score for the latter party's estimates). According to this analysis, supervision decreases the poll estimate error. In contrast, the length of the interview (which had an average time of 12.5 minutes, and a median of 11 minutes, with a few interviews that lasted more than 30 and 40 minutes) does not contribute significantly to the poll estimates' accuracy. Since the questionnaire only had 35 questions, longer interviews may have reflected some interruptions or factors that made them longer but

not necessarily less accurate. Both the respondents' and the interviewers' age contribute to make interviews that are longer on average: The older the respondent (or the older the interviewer), the longer the interview. Of course, the longest interviews are those among older respondents and older interviewers, with an average duration of 14.6 minutes in this study, two more minutes on average than the overall poll average, or 16% longer.

Finally, our analysis does not provide any evidence that "last minute" changes took place or were significant in the 2011 State of Mexico election, and therefore they did not contribute to the poll's estimation error. This does not mean that last minute effects do not take place. Perhaps they are more likely to contribute to the polls' inaccuracy in more contested elections (the State of Mexico election had a 42 point difference between the first and second place). The 2010 state elections, for example, were more contested races, and last minutes changes may have contributed to the poll errors, as argued by Estévez (2010). But we do not have a way to prove it in this article, so this could be a task for future research, obtaining some of the polls conducted by polling firms that year (if they offer a way to have a proxy indicator of last minute changes), or designing an exit poll that measures those effects for future elections.

From the analysis shown in Table 5, we can argue that no single factor explains the inaccuracy of poll results, but a combination of factors that include sampling, interviewer biases and supervision, the context of an interview, and Spiral-of-Silence type of biases. Overall, interviewer effects (especially the sex and age of the interviewer) and interviewer supervision were among the chief explanatory factors of accuracy in the State of Mexico poll that we analyzed. In contrasts, spiral of silence and last minute change effects were rather moderate, and so were the sampling and questionnaire design. It is true that our analysis is generally weak

in its empirical foundations (as evidenced not only by the low R squares, but also by the fact that it is based on one poll in one election that was hardly contested). Still, it is a first formal step in what we hope to be a broader and more collective search for the main sources of polling error in Mexico.

Final discussion

What are the main conclusions from this analysis and in which direction should further research go? The purpose of this article is to start a discussion and continue the analyses of possible sources of errors in pre-election surveys in Mexico. As a young democracy this type of studies might also help us understand issues of pre-electoral polls in new democracies. Based on the empirical evidence, limited to one local election, we can point out to a mix of factors that affect poll accuracy: some interviewer effects, contextual factors, as well as spiral of silence variables. In our attempt to test more hypotheses, further research should also focus on sampling effects, alternative methods to ask the voting question, alternative methods of data gathering, interviewer effects, supervision effects, contextual effects (crime and unsafe polling points), spiral of silence effects, and screening of likely voters.

In order to expand this research it would be convenient to draw evidence from a set of pre-electoral polls conducted in different campaigns or one survey that draws representative samples from electoral areas with different levels of competition. This would allow us to compare the effect of different factors that might affect the measurement of vote choice at different levels of competitiveness. This set of polls might include different experimental designs to measure the set of causes we found significant in this study to explain the discrepancy between the pre-electoral polls and the final election results. Finally, an exit poll could help us assess the

amount of people, who in different contexts decides the direction of their vote within two or one week before the election.

Our findings give us some hints about the estimation errors of 2010, but that collective experience is still in need for a more complete explanation. In addition to the methods and possible effects that we have proposed, there are also other possibilities that may complement the overall picture. For example, Traugott and Wlezien (2011) argue that it has been difficult in the United States to estimate support for what they call "insurgent candidates". In the states where the PRI lost in 2010 against a left-right coalition, the candidate was actually a popular former PRI member that broke away from the governing party as a result of different disputes. It would be interesting to discuss whether this could apply as an "insurgent" effect and whether such category is found useful in Mexico. If yes, it is possible that an insurgent candidate may in fact reflect spiral of silence effects. Another possibility raised by Traugott and Wlezien as a source for estimation error is that it is very difficult to capture "momentum". This elusive phenomenon can take place in different moments of a campaign and it can certainly build up as the election approaches. The presence of momentum at the end of a campaign may be different to last minutes changes, since there may be a trend building up from before instead of just a spontaneous change.

The polling activity in Mexico has evolved with firm steps and setbacks. The estimation errors observed in 2010 certainly damaged the profession's credibility. However, polls remain a vital element of electoral life in the country. We believe that polls should be strengthened, as they serve a noble role of informing the citizens. We hope that the analyses and results discussed in this article serve the polling profession to revise its methodologies and reassess their work, not because they have expired or become useless, but because the object of study they try to capture is a dynamic one, always changing and reflecting new realities.

References

American National Election Studies, ANES (2007), *Time of presidential election vote decision1948–2004*, retrieved September 8, 2011, from http://www.electionstudies.org/nesguide/toptable/tab9a_3.htm.

Basáñez, Miguel (1995), "Public Opinion Polling in Mexico", In Peter H. Smith (ed.), *Latin America In Comparative Perspective: New Approaches To Methods And Analysis*. Boulder, CO: Westview Press.

Brambor, Thomas, William R. Clark, and Matt Golder (2006), "Understanding Interaction Models: Improving Empirical Analyses", *Political Analysis,* vol. 14 no. 1:63-82.

Kam, Cindy K. and Robert J. Franzese, Jr. (2007), *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor: The University of Michigan Press.

Magalhães, Pedro (2005), "Pre-Election Polls in Portugal: Accuracy, Bias, and Sources of Error, 1991-2004", *International Journal of Public Opinion Research*, vol. 17, no. 4: 399-421.

Moreno, Alejandro (2009a), "Encuestas y elecciones en México: La precisión de estimaciones preelectorales en un contexto de cambio" [*Polls and elections in Mexico: The accuracy of pre-election estimates in a changing context*], paper presented at the Second WAPOR Latin American Congress, Lima, Peru, 22-24 April.

Moreno, Alejandro (2009b), *La decisión electoral: votantes, partidos y democracia en México* [*Electoral Choice: Voters, Parties and Democracy in Mexico*], Mexico City, Miguel Angel Porrua.

Nir, Lilach and James Druckman (2008), "Campaign Mixed Message Flows and Timing of Vote Decision", *International Journal of Public Opinion*, vol. 20, nr. 3: 326-346.

Noelle-Neumann, Elisabeth (1993 [1984]), *The Spiral of Silence: Public Opinion, Our Social Skin*, 2nd edition, Chicago, University of Chicago Press.

Przeworski, Adam (1991), *Democracy and the Market: Political and Economic Reforms in Eastern Europe and Latin America*, Cambridge, Cambridge University Press.

Reuters, 17 April 2007, Poll: Nearly half of French voters undecided. [URL: http://www.cnn.com/2007/WORLD/europe/04/09/france.election. poll.reut/index.html.]

Romero, Vidal, and Carlo Varela (2011), "La precisión de las encuestas preelectorales" [*Pree-election poll accuracy (in Mexico)*], *Última Instancia: Revista de Estudios Jurídico Electorales*, Vol. 2, No. 0, Summer. Pp: 30-37.

Romero, Vidal (Forthcoming), "Notas para la evaluación de las encuestas preelectorales: Las elecciones para gobernador de 2010 en México", *Política y Gobierno*.

Sartori, Giovanni (1976), *Parties and Party Systems*, Cambridge, Cambridge University Press.

Schmitt-Beck, R. and T. Faas (2006), "The campaign and its dynamics at the 2005 German general election", *German Politics*, nr. 15: 393–419.

Scheufele, Dietram A. (2008), "Spiral of Silence Theory", in Wolfgang Donsbach y Michael W. Traugott (eds.), *The Sage Handbook of Public Opinion Research*. London, Sage.

Traugott, Michael (2005), "The Accuracy of the National Preelection Polls in the 2004 Presidential Election", *Public Opinion Quarterly*, vol. 69, no. 5, special issue: 642-54.

Traugott, Michael, and Christopher Wlezien (2011), "Media Coverage as a Contextual Explanation for Estimation Errors in Pre-Primary Polls in the United States", paper presented at the 64th Annual Conference of the World Association for Public Opinion Research, WAPOR, Amsterdam, The Netherlands, September 21-23.

Weisberg, Herbert (2008), "The Methodological Strengths and Weaknesses of Survey Research", in Wolfgang Donsbach and Michael Traugott (eds.), *The Sage Handbook of Public Opinion Research*, London, Sage.

Zaller, John (2004), "Floating voters in U.S. Presidential elections, 1948-2000", in P. M. Sniderman and W. E. Saris (eds.), *Studies in Public Opinion*, Princeton, NJ, Princeton University Press.

Appendix

Table A.1. Testing interaction effects for Spiral of Silence hypotheses: OLS regression

	Total absolute error		*Absolute error for PRI*		*Absolute error for PAN*		*Absolute error for PRD*	
	t	Sig.	t	Sig.	t	Sig.	t	Sig.
Sampling effects								
Substitute sample	0.87		0.49		0.48		1.48	
Rural sample	-2.38	*	-2.87	**	1.45		-4.57	***
Questionnaire design								
Voting question at the beginning	-0.36		-0.04		-0.62		-0.08	
Spiral of silence effects								
Minority view in municipality	0.25		0.04		2.19	*	-1.89	
Minority view municipality + Age	-0.36		-0.35		-1.56		1.33	
Minority view municipality + Female	-0.00		0.06		-0.53		0.52	
Minority view municipality + Female + Age	0.18		0.20		0.39		-0.24	
Minority view in the state	-0.38		-0.58		-1.20		1.19	
Minority view state + Age	0.08		0.52		0.56		-1.25	
Minority view state + Female	1.29		1.17		0.89		0.95	

Minority view state + Female + Age	0.99		-0.97		-0.51		-0.80	
Pressure interview	3.11	**	3.42	**	1.40		2.13	*
Contextual effects								
Unsafe polling point	-1.31		-0.29		-3.48	***	0.68	
Interview(er) effects								
Female interviewer	-4.50	***	-5.15	***	-5.74	***	1.51	
Age of interviewer	-1.01		-4.04	***	3.03	**	-0.39	
Experience of interviewer	-0.78		-0.25		0.50		-2.69	**
Supervised interview	-3.11	**	-3.01	**	0.79		-5.44	***
Length of interview	0.75		0.32		0.17		1.51	
Last minute change effects								
	Total absolute error		*Absolute error for PRI*		*Absolute error for PAN*		*Absolute error for PRD*	
	t	Sig.	t	Sig.	t	Sig.	t	Sig.
Undecided	-1.33		-0.97		-1.60		-0.54	
Respondent's age	0.79		-0.08		2.36	*	-0.35	
Female respondent	-0.28		-0.33		0.31		-0.66	
Respondent's age + Female	0.08		-0.07		-0.56		0.34	
(Constant)	8.33	***	17.62	***	3.62	***	9.05	***
ADJ R-SQ	.04		.06		.05		.05	

Significance levels: * p<.05; ** p<.01; *** p<.0