Uncertainty and Errors in the Mexican Elections of July, 2006

W. Luis Mochán
Centro de Ciencias Fsicas, UNAM
A.P. 48-3, 62251 Cuernavaca, Morelos, Mxico

September 7, 2006

Abstract

The information given by the Federal Electoral Institute during and after the elections of July 2, 2006, show errors so large that it is imposible to define a winner with certainty.

1 Introduccin

Every measurement has an associated uncertainty . For example, it is impossible to distinguish two microscopic objects separated a distance of only a few millionths of an inch if we employ a conventional optical microscope, regardless of its quality of perfection. This is a consequence of the resolution limit enunciated by Ernst Abbe in 1973 based on the fundamental properties of light, of wave motion and of its mathematical properties. The fundamental limitations of some measurement processes were recognized by Heisemberg, who stated in 1927 his well known *uncertainty principle*.

In contrast to other measurement processes, counting is essentially exact. Nevertheless, to obtain an exact count a total absence of errors would be required. In a process as complex as a federal election, whose count employs hundreds of thousands or even millions of participants, most of them ordinary citizens, it is virtually impossible to avoid errors. The best trained and honest clerical workers make ocassional mistakes. Thus, the availability of elements that permit an evaluation of the magnitude of the actual errors is of paramount importance. If the errors are substantially smaller than the difference in the number of votes received by the candidates, we could simply ignore them. Nevertheless, if the errors are similar or larger than that difference, resolving it would be an impossible task without a careful recount, in which the main error sources are identified and eliminated as completely as possible.

> *The certainty about all of the emitted votes and their sense becomes relevant in democratic elections to determine the elected candidate, as it is in the best interest of the participanting parties and of society in full that it becomes certain that the vote count was adequately computed, and that*

> *the majority decision is indeed that apparent from the start or else, that the possibility of errors in the count correponding to some electoral booths could lead, after a verification or a recount in the terms specified by the law, to a different result.*
> (Resolution SUP-JIN-212-2006-Inc2 and others, dictated by the Federal Electoral Court of the Judicial Power (TEPJF[1]) on August 5, 2006.)

Fortunately, among the electoral data that the Federal Electoral Institute (IFE[2]) gathers and has made publicly available, there are several that are redundant, that is, data that are not mutually independent and that therefore must obey a set of constrictions, although the data are obtained independently of each other. To mention one example, lets consider the number of votes in one box. It could be determined by adding the number of votes that went to each candidate, the number of votes for non registered candidates and the number of null votes. Even more simply, it could be obtained by counting the number of ballots found in the box. An alternative would be to count the number of actual citizens that went to the polling station to vote, each of which should have deposited one ballot within the box. The number of voters could actually be determined from the number of stamps in the registered voter list, as the name of each voter should have been stamped before he left the station. The three methods should yield the same result. In a few exceptional ocassions, one would expect they wouldn't; maybe a voter took away a ballot, so that the number of ballots in the box would differ from the number of stamps. Otherwise, a voter could make a mistake and deposit his ballot in a different box, so that there would be a box with more and a box with less votes than expected. Another source of errors is an arithmetic mistake when summing the results, or a transcription error during the data capture. Other sources of error include ballot, boxes and/or certificate manipulations.

> *The comparison of these elements allows the verification of the total number of votes*
> (Resolution SUP-JIN-212-2006-Inc2 and others, dictated by the TEPJF on August 5, 2006).

Redundancy is indispensable to judge electoral results. In other fields of human activity we employ redundance continuously. For example, our facial expression and our voice intonation shows happiness when we communicate exceedingly good news, or sadness when we news are bad. If our expression or intonation were inconsistent with our message, our interlocutor would guess that he/she misunderstood and would ask for clarification. Similarly, the communication among computers requires redundancy to detect and sometimes repair the mistakes induced by electromagnetic noise during an electronic transaction.

In order to detect and quantify errors during an electoral count, the certificates filled at the end at each polling station contain redundant fields which should be captured by the electoral clerks and then captured in centers of data gathering and transmission. The clerks, chosen from the citizenship by the Federal Electora authorities before each

[1] From its initials in Spanish.
[2] From its initials in Spanish

election, take a tutorial that presumably trains them to fill these fields, beyond the usual fields such as the number of votes received by each candidate.

In the present paper we analyze some of the errors that are evident in the information that IFE has made publicly available. Using the results of this analysis, we conclude that the expected uncertainty in the electoral results are much larger than the maximum error that is allowable under such a close election.[3] Thus, it becomes paramount to make a full recount of the votes in such a way as to diminish significantly the uncertainty. Otherwise, the legal and the technical requisite of *certainty* would be violated.[4]

> *Certainty is the clear, positive and firm conviction of truth; the absence of doubt about an event or a thing.*
> (Resolution SUP-JIN-212-2006-Inc2 and others, dictated by the Federal Electoral Court of the Judicial Power (TEPJF) on August 5, 2006.)

The paper is organized as follows. In section 2 the information made available in real time during the election night is analyzed. Errors in the data reporting are described: data was manually added, subtracted, and modified, all of which shows that the computer system at IFE is somewhat vulnerable. In section 3 the data base corresponding to the preliminary electoral results program (PREP[5]) is analyzed and consistency checks are applied using to that end the redundant fields. It is shown that, even disregarding those certificates that were not accounted for in PREP for different reasons, in more than 20% percent of the remaining ones it is impossible to apply one or another of the consistency checks as not all relevant fields were captured. Each of the test performed fails in about 40% of the verifiable records, and the number of inconsistencies is of the order of *millions of ballots*. These numbers are only slightly modified if the test is applied to full electoral sections instead of individual ballot stations.[6] In section 4 we analyze the data base corresponding to the District Count, in which some errors are detected. Unfortunately, redundant fields were eliminated from the data base, so that it becomes impossible to apply consistency checks. Nevertheless, a comparison to the PREP data base shows such a small number of changes that it may be safely stated that most inconsistencies went through uncorrected. On the other hand, the observed changes are inconsistent with the hypothesis of simple unbiased error corrections. Finally, the conclusions are presented in section 5.

2 The Night of the Election

The night of the election the IFE gave announced partial results of the Preliminary Electoral Results (PREP) electronically through web page in different computer servers. Using a small program, I gathered copies of the pages published in the site [1] and

[3] The winner had a lead of only 230,000 votes.

[4] As I performed this translation from the original in Spanish, I learned that the Court has made their final veredict on this election, so the criteria of *certainty* was indeed violated.

[5] From its initials in Spanish

[6] Sometimes there were several stations a few feet apart from each other, which could have confused the voters.

# actas procesadas	Votes per booth			Time
	FC	RM	AMLO	
127,710	50	48	47	12:27
127,713	**1,825**	**6,657**	**1,216**	13:50
127,724	115	60	115	13:57
127,732	**-605**	**-2,416**	**-501**	12:33
127,752	378	**1,032**	328	14:03
127,772	**-167**	**-875**	**-219**	12:39

Table 1: Number of votes per booth according to the PREP reports published immediately after the election. The first column shows the total number of processed certificates, columns 2-4 show the average number of votes obtained by some candidates per certificate, and the last column shows the time at which the report was prepared.

I stored them in my computer. Alfonso Baqueiro made a similar program and send copies of the captured pages. Both sets are available in [2] and [3].

The data obtained from these pages were used for the initial phase (stage) of analysis of the presidential election, which may be consulted in [4]. In the present document I will refer solely to the evident mistakes I noticed as I analyzed the average number of votes obtained by each party in each certificate. This number was obtained as follows:

1. With a series of simple programs, I extracted the relevant data from the html code of each of the captured pages.

2. I prepared a file [5] with all the data sorted according to the total number of processed certificates.

3. For each two consecutive records, I subtracted the accumulated votes for each candidate and I divided the result between the difference of the total number of processed booths.

In figure 1 I display the results. Besides de voting behavior received by each candidate and certain pecularities [4] that require an explanation, a series of evident errors appear at the end of the process, when close to 128,000 booths had been processed. The graph shows a series of violent oscillations, so large that get they go beyond the bounds of the plot, reaching values **greater than 6,000**, and on the other hand, **taking negative values smaller then -1,000**. Such results, evidently erroneous, *were a consequence of manipulation in the report of the results of PREP* as shown below. In table 1 I display some of the data obtained on July 3 directly from the PREP web pages [1]. According to the first two lines, in each of the three certificates received after 127,710 had been processed, Felipe Caldern (FC) obtained an average of 1,825 votes, Roberto Madrazo (RM) 6,657 votes and Andrs Manuel Lpez Obrador (AMLO) 1,216 votes. In a similar way, the third and fourth line show that in the eight certificates received after processing 127,724 booths, FC received an average of 605 negative votes, RM 605 negative votes and AMLO 501 negative votes. These numbers, as well as others marked with bold face letters on the table and others not shown by the table are absurd. Nevertheless, there is a simple explanation. Observing the last column, which

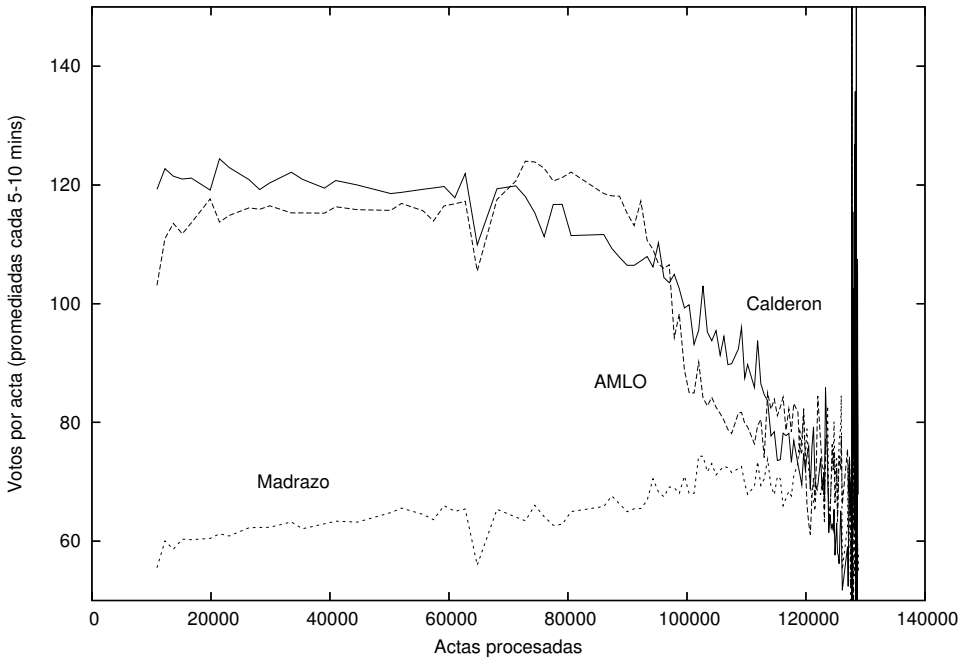
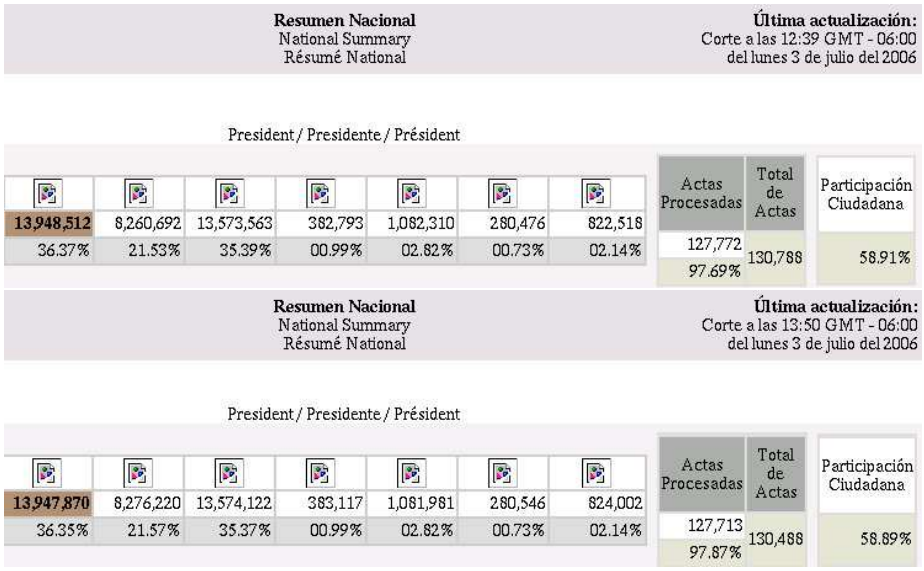


Figure 1: Average votes obtained by each candidate per booth as a function of the number of booths reported by PREP on the night of the election and on the next day. The average is done over all the booths processed between two consecutive records from my data base, corresponding aproximately to five minute intervals.



Resumen Nacional
National Summary
Résumé National

Última actualización:
Corte a las 12:39 GMT - 06:00
del lunes 3 de julio del 2006

President / Presidente / Président

							Actas Procesadas	Total de Actas	Participación Ciudadana
13,948,512	8,260,692	13,573,563	382,793	1,082,310	280,476	822,518	127,772	130,788	58.91%
36.37%	21.53%	35.39%	00.99%	02.82%	00.73%	02.14%	97.69%		

Resumen Nacional
National Summary
Résumé National

Última actualización:
Corte a las 13:50 GMT - 06:00
del lunes 3 de julio del 2006

President / Presidente / Président

							Actas Procesadas	Total de Actas	Participación Ciudadana
13,947,870	8,276,220	13,574,122	383,117	1,081,981	280,546	824,002	127,713	130,488	58.89%
36.35%	21.57%	35.37%	00.99%	02.82%	00.73%	02.14%	97.87%		

Figure 2: PREP reports obtained in *real* time, showing a drop in the number of votes and in the total number of processed certificates.

shows the time at which each report was prepared, wa can tell thet the lines are not arranged in chronological order. For example, at 12:39, 127,772 booths had been processed and one hour later, at 13:50, 127,713 booths were processed, which shows that not less than 59 certificates that had previously been processed were eliminated from the reports. In figure 2 the coresponding PREP reports are shown. Apparently this error was related with the elimination of the votes from abroad, which later on were returned back, generating another series of errors. Even correcting the modified certificates with the incorporation of the votes from abroad [6] there are mistakes remaining in the report of the data.

It is important to remark that the anomalies pointed out in this section are mistakes in the PREP *report*, not in the data capture. Therefore, they have no direct consequences in the results of the election. As a matter of fact, these errors could be considered harmless. To err is human, as well as admitting, correcting and explaining them would have been a sign of decency. Instead of that a very intense propaganda campaign was initiated to implant through endless repetition the notion that all aspects of the election process were impeccable and perfect. In the presence of this propaganda, it is important to emphasize that

1. The PREP presented obvious errors as documented above.

2. These errors show that officials at IFE have the capacity of iterfering with the computers that made this reports, adding, eliminating and modifying data.

3. The computer system in IFE, or at least that part in charge of reporting the results from PREP, is not robust and it may be interfered.

4. Tha absence of an explanation of these anomalies and the enormous propaganda pretending to induce the notion of a perfect process can not but produce distrust about the other stages in the election process.

3 PREP Data Base

A few days after the election, IFE made available a series of data bases reporting the PREP results. Those data base were available at different web sites. The one I used at first was from [7], where it is not available anymore. Yet, in [8] I mounted a copy. Afterwards, more complete versions (with additional fields) of the same data base were made available [9] at IFE's official website [10].

From an inspection of the certificates reported in the PREP data base for the presidential election [9] it is inferred that

1. It contains 117,287 records.

2. 13,201 records are missing and would be required to account for the 130,488 electoral stations installed (not counting those corresponding to the votes from abroad) and another 300 records in order to include the results from the voting abroad. There are many reasons that explain this absences, the main one of which is the presence of inconsistencies that prevented incorporating those certificates in the PREP.

3. Among the records that are present in the data base, there are 24,148 that are incomplete. In these there are 31,302 numerical fields that were left empty. In table 2 the absent fields and the number of times that were left empty is shown.

4. The fields that were found to be absent in the certificates that were accounted for in the PREP count are those that are apparently unrelated to the election results, as the don't involve the votes obtained by any of the registered candidates. The records whose missing fields were directly related to the electoral results were registered in another data base and were not accounted for in the PREP. Nevertheless, it is important to remark that those fields that are frequently missing from the records that were deemed acceptable for the PREP were precisely those that *provide the redundance that allow us to verify if there were errors or misdeeds in the corresponding electoral booth*. The absence of those **makes it impossible to apply the consistency checks designed to detect errors and/or manipulation of the electoral results.** Thus, there is no way to find out if the data in those records is correct, is mistaken or is fabricated.

5. **Thus, there are 24,148 records that ammount to 21% of the number of certificates accounted for in the PREP and 18% of the total amount of electoral boxes in the presidential election for which the consistency tests may not be applied.**

6. In 8,153 of those (6% of the total) it is impossible to know whether the number of ballots deposited in the box is greater or smaller than the number of blank

Meaning	Name of the field	Empty
Number of ballots found in the box	NUM_BOLETAS_DEPOSITADAS	7,637
Number of ballots received before the installation of the booth	NUM_BOLETAS_RECIBIDAS	753
Number of blank ballots after the booth closed	NUM_BOLETAS_SOBRANTES	1,378
Votes received by unregistered candidates	NUM_VOTOS_CAN_NREG	12,997
Null votes	NUM_VOTOS_NULOS	4,481
Number of voters, according the the stamps found in the voter's list	TOTAL_CIUDADANOS_VOTARON	4,056.
Ballot related: Number of records with for which of one or more of the ballot related fields is missing, such as NUM_BOLETAS...		8,153
Vote related: Number of records for which the number of voters, null votes or votes for non-registered candidates is missing		19,457
Ballots and voter related: Number of records for which the total number of ballots deposited in the box or the number of voters is missing		9,862
Empty fields		31,302
Incomplete records		24,148

Table 2: Empty fields in the data base of certificates that were accounted for in the PREP count, numer of times that each field was absent and meaning of the field.

ballots before the election minus the number of blank ballots remaining after the election. Thus, we cannot find out if there were voters that took away *illegally* the ballots that should have been deposited in the boxes, or if ballots were taken away or added to the boxes irregularly.

7. In 19,497 records (15%) it is impossible to ascertain whether the sum of the all the counted votes[7] agrees with the number of voters, as the later number, the number of null votes and/or the number of votes for unregistared candidates is missing. Thus, the total number of votes is uncertain.

8. In 9,862 records (8%) it is impossible to know whether the number of ballots in the box correspond to the number of voters, since one of the corresponding fields is absent.

9. In 22,147 records (19% of the total) it is impossible to know if the number of counted votes correspond to the number of ballots in the box, since the last data, the number of null votesand/or the number of votes for unregistered candidates was not recorded.

10. In some records it is impossible to apply more than one of the previous tests. Thus, the sum (8,153+19,497+9,862+22,147=59,659) of records with different types of absent fields is more than the total (24,148) of incomplete records.

[7] Including null votes

11. Of the 109,134 records where the number of ballots in the box may be compared to the number of received blank ballots and the number of remaining ballots is possible, there are 17,465 (16%) in which the number of deposited ballots is greater by 788,077 when compared to the difference between the number of received and remaining ballots. That is, an average of 45 ballots in excess are found in each of the corresponding boxes. There are also 32,758 records (30%) where the number of deposited ballots is smaller by 716,489 as compared to the difference between received and remaining ballots. That is, there is an average of 22 ballots missing from each of these boxes. **There is a total of 50,223 booths (46%) with this type of error, which involve 1,504,566 ballots**. Subtracting the missing ballots from the ballots in excess we obtain a net excess of 71,588 ballots.

12. Of the 97,790 records where the number of counted votes may be compared to the number of voting citizens, there are 22,419 (23%) in which the number of votes surpass by 719,857 the number of voters. That is, in average, a surplus of 32 votes in each of these booths. There are also 22,391 records (23%) in which the number of counted votes is smaller by 1,043,907 than the number of voters. That is, 47 missing votes on the average for each of the corresponding booths. **This ammounts to 44,810 records (46%) with this kind of error, involving 1,763,764 votes.** Subtracting missing from excess votes we obtain a a net defect of 324,050 votos.

13. Of the 107,425 records where the number of deposited ballots may be checked against the number of voters, there are 17,681 (16%) where the number of ballots surpasses the number of voters by 876,422. That is, within each of the coresopnding boxes there are an average of 50 ballots in excess. There are also 26,342 records (25%) in which the number of ballots deposited is 1,474,589 below the number of voters. That is, on the average there are 56 ballots missing from each of the corresponding boxes. **Thus there are 44,023 records (41%) with this kind of inconsistency, involving 2,351,011 ballots.** Subtracting missing ballots from those in excess we obtain a deficiency of 589,167.

14. Of the 95,140 records which allow a comparison between the total number of counted votes to the number of deposited ballots, there are 17,889 (19%) in which the number of votes is larger in 457,415 than the number of ballots. That is, 26 votos were counted on the average without finding a corresponding ballot within the electoral box. There are also 9,357 records (10%) for which the number of votes is lower by 230,927 than the number of deposited ballots. That is, an average of 25 ballots were skipped and the corresponding votes were not counted for each of the coresonding boxes. **There are 27,256 records (29%) withi this kind of inconsistency involving 688,342 votes.** Subtracting the missing votes from the extra votes we obtain a net count of 226,448 votes above the number of ballots.

The results above are summarized in table 3. As it is clearly seen, *each of the verifiable inconsistencies involve of the order of hundreds of thousands or even millions of ballots*.

Test	Verifiable records	Recs. where larger	Size.	Recs. where. smaller	Size.
Deposited ballots vs. received - remaining	109,134	17,465 (16%)	788,077	32,758 (30%)	716,489
Votes vs. voters	97,790	22,419 (23%)	719,857	22,391 (23%)	1,043,907
Deposited ballots vs. voters	107,425	17,681 (16%)	876,422	26,342 (25%)	1,474,589
Votes vs. deposited ballots	95,140	17,889 (19%)	457,415	9,357 (10%)	230,927

Table 3: Results of several tests applied to the PREP data base. The test is briefly described, the number of records containing enough information to apply the test, the number of inconsistent records with one or the other sign and the size (number of votes, ballots, voters) involved in the inconsistencies.

There are some tentative explanations for some of the inconsistencies described above and summarized in table 3 which ought to be explored. For instance, it is not unconceivable that an important number of citizens became confused in those sections[8] which contained one or more booths adjacent to a main booth[9] and deposited their ballot in the box corresponding to the wrong booth. This aparently harmless mistake (whatever is missing from one box would appear in a nearby box) may be eliminated from the inconsistency checks if the test are not performed individually at the booth level, but aggregated at the section level. Thus, the possible confusion mentioned above with be eliminated when the data of a given booth is added to those of the adjacent booths. Any remaining error would require alternative explanations.

1. In the PREP data base for counted booths there is data about 59,084 sections. I applied the previous tests to those sections all of whose records had enough data, i.e., if a single record within a section is not verifiable, then the whole section is not verifiable. Thus the information that follows refers to a somewhat smaller number of booths than that analyzed above.

2. Of the 51,538 sections for which which the number of deposited ballots may be compared to the difference between the number of received and remaining blank ballots, there are 8,299 (16%) in which the number of deposited ballots is larger by 632,682. On the average, there are 76 extra deposited ballots for each of these sections. There are also 19,117 sections (37%) for which the number of deposited ballots is smaller by 580,875. Tha is, on the average 30 ballots were not deposited in the boxes of each of these sections. **There are 27,416 sections (53%) with this kind of error involving 1,213,557 ballots.** Subtracting the missing ballots from the extra ballots, a net excess of 51,807 ballots is found.

3. Of the 42,093 sections in which the number of counted votes may be compared to the number of voters, in 11,209 (27%) it is larger by 517,866. Thus, there are 46 extra votes on the average for each of these section. There are also 11,

[8] Each state was divided in electoral districts consisting of many sections, each of which had one or more polling stations.

[9] As a maximum number of votes was established for any individual booth, several booths were assigned to sections with many voters. They were frequently located very close to each other; sometimes only a few feet apart.

Test	Verifiable sections	Secs. where larger	Size.	Secs. where smaller	Size.
Deposited ballots vs. received - remaining	51,538	8,299 (16%)	632,682	19,117 (37%)	580,875
Votes vs. voters	42,093	11,209 (27%)	517,866	11,289 (27%)	761,954
Deposited ballots vs. voters	50,035	9,312 (19%)	685,298	15,838 (32%)	1,213,921
Votes vs. deposited ballots	40,057	11,039 (28%)	345,112	5,508 (14%)	156,094

Table 4: Results of several tests applied to the PREP database PREP aggregating data from all booths wihin a section before applying each test. The tests are briefly described, the number of sections with enough information to apply the test, the number of sections with one or another kind of inconsistency and the size (number of votes, ballots, voters) involved in the inconsistency.

289 (27%) sections in which the number of votes is smaller by 761,954 than the number of voters. That is, there are 67 votes missing on the average from each of these stations. **There are 22,498 sections (53%) with this kind of error, involving 1,279,820 votes**. Subtracting the missing votes from the extra votes we obtain a net defect of 244,088.

4. Of the 50,035 sections in which it is possible to compare the number of deposited ballots to the number of voters, there are 9,312 (19%) in which the number of ballots is larger by 685,298. That is, there are an extra 74 ballots deposited in the boxes of these sections. There are also 15,838 sections (32%) in which the number of vallots is smaller by 1,213,921 than the number of voters. That is, 77 ballots are missing on the average from the boxes of each of these sections. **There are 25,150 sections (50%) presenting this kind of inconsistencies, which involves 1,899,219 ballots.** Subtracting missing from extra ballots we obtain a defect of 528,623 ballots.

5. Of the 40,057 sections in which the number of counted votes may be compared to the number of deposited ballots, there are 11,039 (28%) in which the number of votes is larger by 345,112. Tha means, ther are 31 votes that were counted on the average on each of these sections, although they didn't correspond to a ballot deposited in the corresonding boxes. There are also 5,508 sections (14%) in which the number of deposited ballots is larger by 156,094 than the number of votes. That is, there were on the average 28 deposited ballots whose corresponding votes were not counted in each of these sections. **There are 16,547 sections (41%) that show this type of inconsistency involving 501,206 votes.** Subtracting missing from extra votes, we obtain a net excess of 189,018 votes above the number of deposited ballots.

The results above are summarized in table 4 which shows than even after aggregating the data section-wise, errors and inconsistencies survive, marginally decreasing their size. Nevertheless, they still involve hundreds of thousands and even millions of ballots, votes, voters, etc. It should be emphasized that the numbers above are *lower bounds to the actual number of inconsistencies*, as there were many records and sections for which the tests could not be applied.

There are many other anomalies in the data bases of PREP, such as 127 non-special[10] booths for which the number of voters that were signed in[11] exceeds by more than 10 the nominal list of voters.[12] Another kind of error becomes evident when each certificate's arrival time (HORA_RECEPCION_CEDAT) at the Centers for Data Aquisition and Transmission (CEDAT[13]). It turns out that the certificates of 68 booths were received *the day before the election*[14] and 1,278 certificates received during the election day but *before closing time*.[15] These[16] and many other errors and anomalies in the PREP are described in [4, 11] and the links there contained.

4 District Tallies

On Wednesday, July 5, 2006, the official vote tally (Computos Distritales (CD's) in Spanish) took place on each of the 300 electoral districts took place. On July 13 I obtained the corresponding data bases published by IFE at [12]. Unfortunately, they omitted all of the redundant fields which would have allowed a consistency check on-their data. For example, they do not mention how the number of ballots deposited in the boxes (which was the field called NUM_BOLETAS_DEPOSITADAS in the PREP data bases), the number of received ballots (NUM_BOLETAS_RECIBIDAS), the number of remaining ballots (NUM_BOLETAS_SOBRANTES), nor the number of voters that were signed in (TOTAL_CIUDADANOS_VOTARON). Thus, **it is impossible to verify whether the inconsistencies described in Sec. 3 were or not corrected during the CD**.

Nevertheless a shallow analysis may be performed by examining the contents of the data base and by comparing its contents to those of the PREP.

1. Even thought there was no strict time constrains for the CD in contrast to PREP, and although it should have resolved all kinds of ambiguities, its results are not flawless, as shown by the presence of 311 records with empty fields (TIPO_ACTA, PAQUETE_ENTREGADO y CASILLA_INSTALADA).[17]

[10] Several so called special booths were set up to allow voters to vote outside of the section in which they were registered, to allow travellers to vote. This booths lacked a list of voters.

[11] The corresponding field in the database was called TOTAL_CIUDADANOS_VOTARON.

[12] Only ten additional votes were permitted to allow the electoral clerks to vote.

[13] From its Spanish initials.

[14] 45 from the state of Guanajuato, 16 from the State of México, 3 from Distrito Federal, 2 from Baja California, 1 from Chiapas, and 1 from Chihuaha

[15] 147 from Edo. de México, 119 from Tlaxcala, 108 from Jalisco, 98 from Guerrero, 92 from Michoacán, 89 from Oaxaca, 84 from Distrito Federal, 73 from Guanajuato, 55 from Hidalgo, 49 from Zacatecas, 36 from Nayarit, 35 from Querétaro, 34 from Campeche, 33 from Tamaulipas, 32 from Sinaloa, 23 from Yucatán, 21 from Puebla, 20 from Coahuila, 16 from Aguascalientes, 15 from Sonora, 14 from Durango, 13 from Tabasco, 12 from Morelos, 11 from Quintana Roo, 9 from Nuevo León, 9 from Chihuahua, 9 from Chiapas, 8 from San Luis Potosí, 5 from Veracruz, 5 from Baja California, 3 from Colima, 1 from Baja California Sur.

[16] Maybe these errors are due to a simple mistake while writting down manually the reception time. Adding a full day to those certificates which were received more than 24 hours before being captured, no records appear to have been actually received before the election and only 45 remain as received before closing time.

[17] Corresponding to Edo. de Mxico (42), DF (27), Veracruz (21), Oaxaca (20), Jalisco (19), Puebla (16), Guanajuato (14), Chiapas (12), Nuevo Len (12), Michoacn (12), Chihuaha (9), Guerrero (9), Tamaulipas (8), Sinaloa (8), Baja California (8), Coahuila (7), Sonora (7), San Luis (7), Hidalgo (7), Tabasco (6), Yucatn

2. Another evidence of negligence is the presence of 20 records (14 from Oaxaca, 2 from Edo. de Mxico, 2 from Guerrero, 1 from Baja California and y 1 from Hidalgo) in which the total number of votes was zero: no votes were registered in favor of any party, alliance or coalition, there were no votes for unregistered candidates and there was no null vote.

3. The data base of the CD for the presidential election contains 13,501 records that were absent from the PREP.

4. The results of those records are remarkably different from the global results. For example, FC reduced his percentage of the vote down to 31.02%, loosing more than 4%, while RM increased his percentage up to 30.86% gaining more than 8%. AMLO remains almost invariant at 35.59%.[18]

5. The records that are present in both the CD and the PREP data bases differ in only 4,151 records.

6. Of these records, there are only 1,243 in which the votes for FC are modified, 1,278 in which the votes for RM were modified and 1,458 that affected AMLO. A statistical analysis [4] of the modifications suffered by each candidate shows that they are incompatible (i.e., probability less than 10^{-17}) with the hypothesis of a simple correction of unbiased accidental mistakes.

7. According to IFE, only 2,873 electoral packages were recounted during the CD.

8. It seems unbelievable that after the PREP produced so many inconsistencies, its revision yielded such a small number of modifications. As all the redundancy that could have allowed a consistencies check of the CD data was eliminated from the official data bases, there is no alternative but to state that almost certainly the number of inconsistencies remaining in the CD are of the same order of magnitude as those in the PREP, that is, there are tens of thousands of booths, more than 50% of the total, in which there are inconsistencies that involve millions of votes.

5 Conclusiones

This work shows that there were errors and manipulation in the computer system in charge or reporting in real time the results of PREP during the election night and the following day.

It also shows a large number of inconsistencies in the PREP certificates, such as they were reported in the data bases that were made publicly available by IFE. In many cases, the number of ballots deposited in the boxes turned out to be inconsistent with

(5), Morelos (5), Zacatecas (4), Quertaro (4), Durango (4), Tlaxcala (3), Quintana Roo (3), Nayarit (3), Aguascalientes (3), Colima (2), Campeche (2) and Baja California Sur (2).

[18] Even more remarkable is the fact that after correcting the numbers above for the fraction of rural and urban populations, the results agree closely with those predicted by most opinion polls before the election, unlike the actual global results.

the number of blank ballots received before the election and the number of blank ballots remaining after the election, the number of counted votes disagreed with the number of citizens that signed in as voters during the election, the number of ballots deposited in the boxes was inconsistent with the number of voters and with the number of counted votes. Each of these errors appears in tens of thousands of booths and involves hundreds of thousands or even millions of votes. The size of the errors, distributed in about half of the more than 130,000 booths, is several times larger than the difference between the number of votes obtained by the main candidates.

It is quite probable that many of the inconsistencies has its origin in simple human errors made without malice. It has been speculated in the Mexican press that most of them originate in the confusion of some citizens due to the closeness between basic and adjacent booths corresponding to the same section. As part of this work I verified that this confusion could indeed produce some of the inconsistencies, but that it is not enough to explain their magnitude, which remains even after aggregating the data section-wise.

Other errors could have been simple clerical errors made when filling the certificates. Probably the process was too complex; probably the training was insufficient. For example, I found about 1,000 records where the number of received ballots coincides exactly with the number of remaining ballots, and about 600 where it coincides with the number of deposited ballots. Most probably, the clerks couldn't differentiate the meaning of the corresonding fields in the certificates. Nevertheless, this kind of confusions is not enough to explain the magnitude of the errors. On the other hand, other explanations which are not as inocent should not be discarded *a priori*.

The evident failures of the PREP should have been corrected routinely during the CD. However, the number of electoral packages which were actually recounted during the CD and the number of records which were modified is much smaller than the number of inconsistencies in the PREP. This is evidence that most of the inconsistencies were not corrected. Unfortunately, the CD data bases do not contain the redundatn fields that could have allowed a verification of the previous statement.

In summary, I have showed evident errors and inconsistencies in several stages of the presidential election of July, 2006. The present analysis has not overlooked the many anomalies found in the electoral results, in their statistical behavior and in the temporal evolution during their campture in the PREP and the CD. These are controversial issues whose study requires some degree of interpration. Many analysis along those lines may be found in [4, 11]. On the other hand, the results presented here may be considered as simple hard data.

Disregarding their origin, the errors and inconsistencies are so large that without eliminating them it would be impossible to designate *with certainty* a winner for the presidential election. With a *measurement* that turned out not to have the necessary resolution, I see no alternative but to make another more accurate measurement. It is indispensable to make a full recount in order to resolve the election.[19]

[19]The electoral court made a final decission on September 5, 2006, declaring FC winner as he obtained about 230,000 more votes than AMLO, his closest competitor. The court knew of some of the inconsistencies described here, but dismissed them as *it was not proved that their absence could have changed the results.*

References

[1] http://www.elecciones2006.unam.mx/PREP2006/index_contenido.html

[2] http://em.fis.unam.mx/public/mochan/elecciones/prep

[3] http://em.fis.unam.mx/public/mochan/elecciones/prep1

[4] http://em.fis.unam.mx/public/mochan/elecciones/

[5] http://em.fis.unam.mx/public/mochan/elecciones/diferenciasporcasilla.dat

[6] Rici Lake, http://em.fis.unam.mx/public/mochan/elecciones/archivos/msg00494.html

[7] http://prep2006.grc.com.mx/extraccion-servlets/presidente.txt.

[8] http://em.fis.unam.mx/public/mochan/elecciones/fullprep.txt

[9] http://www.ife.org.mx/documentos/proceso_2005-2006/prep2006/bd_prep2006/bd_prep2006.htm

[10] http://www.ife.org.mx/

[11] http://analisis.elecciones2006.unam.mx/index.php

[12] http://www.ife.org.mx/documentos/computos2006/bd_computos06.htm