



Report on an Analysis of the Representativeness of the Second Audit Sample, and the Correlation between Petition Signers and the Yes Vote in the Aug. 15, 2004 Presidential Recall Referendum in Venezuela

This study was conducted by The Carter Center and confirmed by the OAS in response to a written request from Sumate presented to The Carter Center Sept. 7, 2004. Sumate asked that The Carter Center evaluate a study performed by Professors Ricardo Hausmann and Roberto Rigobon.

The Hausmann/Rigobon study states the second audit conducted Aug. 18-20 and observed by The Carter Center and the OAS was based on a sample that was not random and representative of the universe of all voting centers using voting machines in the Aug. 15, 2004, recall referendum.¹ The study further indicates that the correlation coefficient (elasticity) for the correlation between the signers and the YES votes for the sample was 10 percent higher than that for the universe. The Hausmann/Rigobon study came to these conclusions through an analysis of the exit poll data, petition signers data, and electoral results data provided by Sumate.

1 Objectives of the Carter Center Study

1. Determine the correlation between the number of signers of the presidential recall petition and the electoral results of the Aug. 15 recall referendum.
2. Compare the characteristics of the universe of voting machine results with those of the sample for the 2nd audit performed Aug. 18.

3. Eo9ih9C ri2 0 0 12 609701904 234.42038(acs of n Tmh tfrTm(Com)Tj12 0 2642 112.64904 234.420

2 Data Sources

The data used to perform this study was officially received from the CNE. The following data was used:

1. The Voters List (REP) used for the Aug. 15 recall referendum officially received from the CNE July 30, 2004.
2. The “*cuadernos de reparo*” database containing the valid signatures and the “repairable” signatures used during the *Reparos* process and officially received from the CNE.
3. The rejected signature list (signatures that could not be repaired during *Reparos*) obtained from the CNE.
4. The electoral results file of voting machines for the Aug. 15, 2004, recall referendum, per machine, officially received from the CNE Aug. 18, 2004.
5. The sample generation program for the Aug. 18 audit including the source code, the executable file, the input file with the universe and the generated sample.

3 Methodology

The four data sources were loaded into different tables² on an IBM DB2 database to facilitate processing. The following calculations were performed:

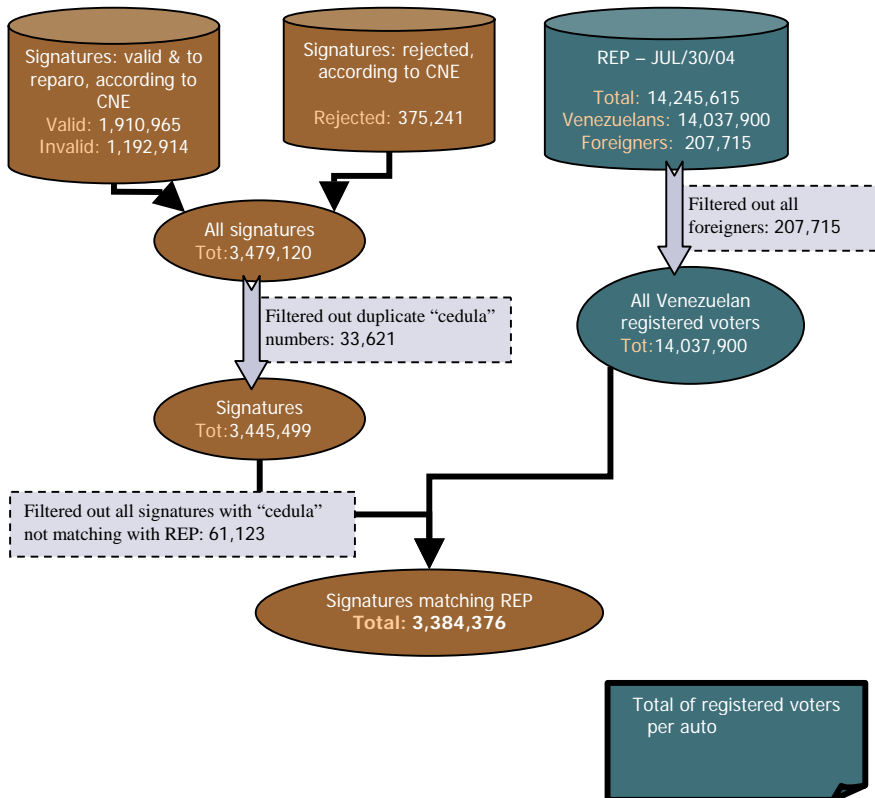
1. The number of voters per voting center was calculated from the REP, excluding foreigners³.
2. A single table of signers in the database was loaded from the *cuadernos de reparos* file and the rejected signatures file, eliminating duplicate ID card numbers.
3. The number of signers per voting center was calculated from the table of all signers by matching the ID card number in the signer’s table with the ID card number in the REP and aggregating the signers into voting centers identified by voting center on the REP table.
4. The YES and NO votes per voting center were calculated by adding the electoral results from each voting machine in that center.
5. A final results table was produced with the following columns for each voting center:
 - a. State
 - b. Municipality
 - c. Parish
 - d. Voting Center Number
 - e. Total registered voters in the voting center
 - f. Total signers registered in the voting center
 - g. Total YES votes
 - h. Total NO Votes

² A table in a relational database is a storage entity where all records have the same columns. A database can have multiple tables and allows operations between tables.

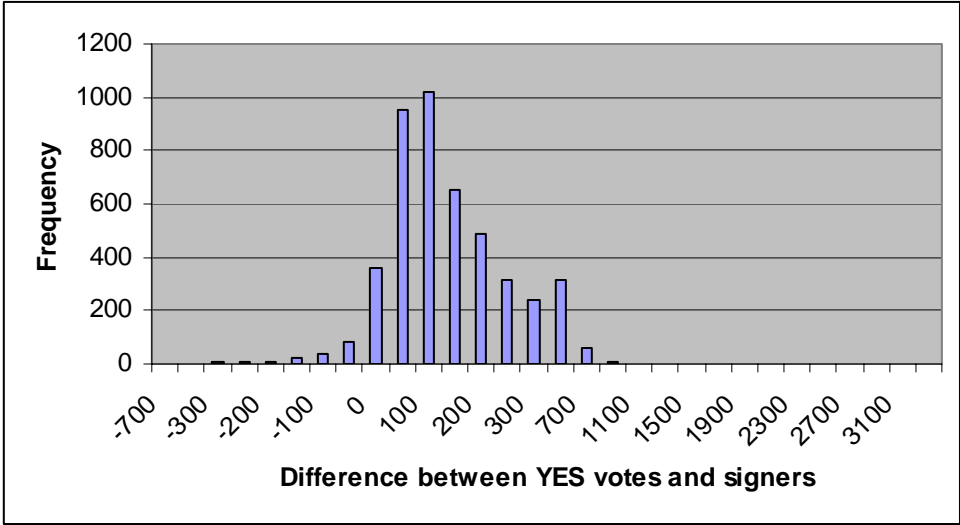
³ Foreigners are not allowed to vote for president in Venezuela, consequently they cannot recall him either.

The final results table was exported into an Excel file. The correlation coefficients were calculated using SPSS version 12.

Additionally an Excel worksheet with only the voting centers that had a mesa (voting station) audited in the Aug. 18 random sample used to perform the second audit was generated to evaluate the representativeness of the sample in the universe.



The sample generation source and executable pr



Universe	0.988
Sample	0.989

MSBDC BLS I

Average appearance of a mesa in a sample	25.05
Standard deviation	5.52
Minimum appearance	14
Maximum appearance	40

5 Conclusions

The sample drawing program used Aug. 18 to generate the 2nd audit sample generated a random sample from the universe of all *mesas* (voting stations) with auto