Forecasting Election Results

by D. Milledge and Mary J. Mills

Summary: This paper describes the methods used on a Ferranti Pegasus Computer to analyse the early results of the 1959 General Election in the United Kingdom and forecast the final state of the parties. Some new methods of data handling and checking are described in detail and the results summarized.

INTRODUCTION

Analysing election results is a "natural" computer application because it is a task demanding rapid and accurate data processing. It is a rather specialized project but it raises several problems of general interest.

The Ferranti Computer Centre in London collaborated with *The Guardian* on election night to analyse the early results and forecast the final state of the parties. The complete job, including the forecast, was run on a Pegasus computer. As an independent check, a Sirius computer calculated the swing for each party and this was used as the basis for a separate forecast.

DATA CHECKING

It was decided from the outset that the usual methods of data checking would not be suitable for dealing with election results, and different methods were devised. These were based on the principle that the computer itself should do most of the comparing and checking of results, because it could do this far more rapidly and accurately than its human operators. Such information as was available was prepared in advance, checked by conventional methods, and stored on magnetic tape. This comprised, for each constituency, the full 1955 results, the 1959 electorate, the parties standing in 1959, the region number, the town number (large towns only), and a marker to show whether or not the previous M.P. was standing again.

The 1959 results were obtained from three sources: B.B.C. television, radio, and telephone from the London office of *The Guardian*. None of these could be regarded as very satisfactory media for data transmission, since all are transitory, and it was decided to feed two of them to the computer and refer discrepancies to the third (television). At each source (except television) three people copied the results and a fourth compared them before sending them for punching. The results from radio and telephone were sent independently to the computer and then compared and checked by it. Differently coloured cards, tapes and teleprinter paper were used for the three streams of information; the organization is illustrated in Fig. 1.

The results were recorded on coloured cards, which held the following information:

Constituency number
Constituency name
Votes for each candidate, listed in a standard party order
Majority.

The constituency numbers used were the Press Association numbers: these were read over the telephone but had to be inserted afterwards on the radio cards.

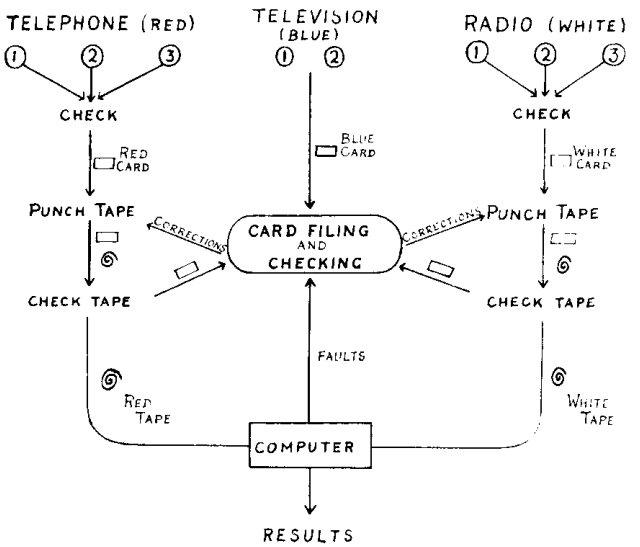

FIG. 1.—The handling of results.

The form of punching was very simple: each result was introduced by → and the constituency number, then followed the votes for each candidate, the majority, and finally an * to terminate the result. For example, the Cornwall North result was punched as shown on the left below:

→148	Constituency number
16701	Conservative votes
3389	Labour votes
15712	Liberal votes
989	Majority.
*	

THE INPUT ROUTINE

As each data tape was read the input routine carried out the normal checks against punching errors such as impermissible characters or layout: in fact there were very few of these as the punching conventions were simple. Further tests were then made to ensure that each result was consistent with information previously recorded for that constituency. Fig. 2 is a simplified flow chart showing the procedure.

After reading each constituency number, the computer printed it out as a record and checked that it was not

greater than 630. The votes for each candidate, and the majority were then read in; if the majority was not known it could be replaced by an oblique stroke. The input routine referred to the appropriate constituency record to discover the party allegiance of each candidate, and to check that the number of candidates was correct. If any of these checks failed the word FAULT was printed, followed by a code number to indicate the

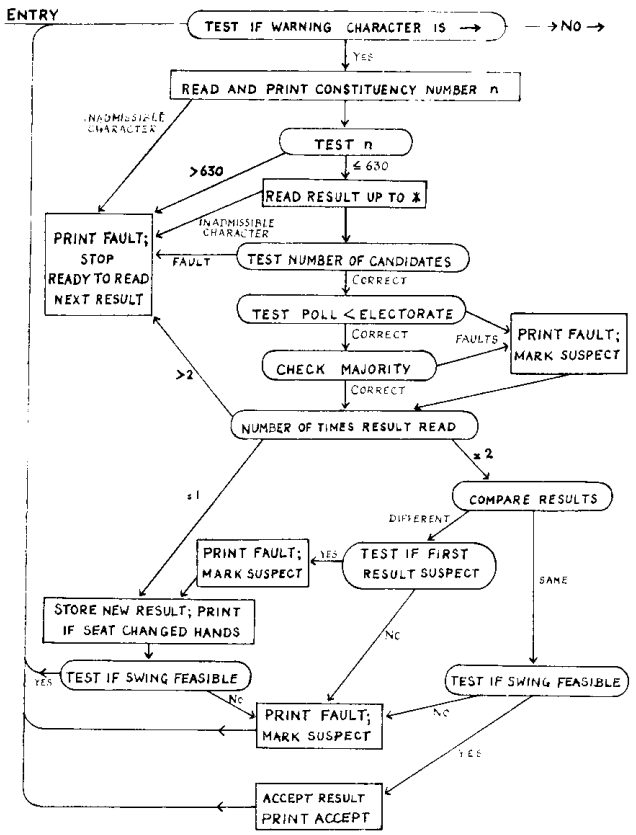

FIG. 2.—The main input routine.

nature of the fault; the computer then stopped ready to re-read the corrected result or to proceed to the next one.

The input routine totalled the votes cast in each constituency and checked that this did not exceed the electorate. It calculated the majority and checked that this agreed with the majority punched on the tape. It also calculated the swing to or from the Conservative party and checked that this lay between $+8\%$ and -10% in straight fights, or between somewhat wider limits in other contests. It was thought that these arbitrary limits would be wide enough to cover all likely results but narrow enough to detect gross errors. If any of these tests failed a FAULT indication was printed, as described in the previous paragraph, but the computer did not stop; instead, it recorded the result but marked it suspect so that it was not used by the main program.

The input routine had to read and compare the radio and telephone results for each constituency. These two results did not normally arrive together: the first was stored but the figures were not used until they were confirmed by the second. If the two results differed, a FAULT indication was printed and the result was marked suspect; if the first result was already suspect it was replaced by the second. If the two results were the same and the swing was feasible, the computer printed ACCEPT, and that result was then made available for use by the main program. A further result for a constituency which had been accepted was treated as a fault.

Besides the main input there were facilities for correcting, confirming or cancelling results. A corrected result had to be punched in full but introduced by $>$ instead of $\rightarrow$; it was subjected to the same checks as an ordinary result but only had to be read once. If a result had been stored as suspect it could be confirmed and accepted by punching v and the constituency number. A result could be cancelled by punching $\neq$ and the constituency number. If a result were known without any details of the poll, it could be read by punching $=$, the constituency number and a code for the winning party: such data was of no use for calculating swings, but it prevented the computer predicting seats which were already known.

The elaborate checking in the input routine was fully justified by the event. Although three people recorded and a fourth checked, both on radio and telephone, the frequency of errors on the cards was about 8%. There were rather less errors on the telephone cards where we had control over the speed of transmission and could identify constituencies by number as well as by name. Despite this high error rate a subsequent check has failed to reveal a single error in the results accepted by the input routine and used for prediction on election night. High accuracy was essential because the prediction was very sensitive to small changes in swing: the average swing, being small, could easily have been upset by a few gross errors. Large errors, such as the transposition of two parties, did occur but were rejected by the input routine.

The data used by the computer was, therefore, accurate, but there were always a number of telephone results awaiting confirmation from the radio and vice versa. Provided they were not suspect, these results were used to show the current state of the parties but not to calculate swings.

PREDICTION METHODS

Most methods of election forecasting, whether applied before or during the election, are based on the simple concept of a "swing." The swing towards a given party may be defined as the increase, from one election to the next, in the percentage of votes cast for that party. In Acton, for example, the Conservative party in 1955 polled 20,120 out of the 40,765 votes cast: that is 49·4%. In 1959 the corresponding figure was 51·2% and the swing to Conservative in this constituency was therefore $(51{\cdot}2 - 49{\cdot}4) = +1{\cdot}8\%$. An average swing for each

party may be estimated either from opinion polls before the event or from the early results. These swings may then be applied to the 1955 figures for each constituency to predict the proportion of votes going to each party, and hence to predict the results for that constituency.

On this occasion the results already declared were used to estimate the swing. For forecasting purposes, constituencies were divided into three groups as follows.

1. Straight Conservative *v.* Socialist contests in both the 1955 and 1959 elections.
2. Three-cornered contests between Conservative, Socialist and Liberal in both elections.
3. Three-cornered contests between Conservative, Socialist and Liberal in 1959, which had been straight Conservative *v.* Socialist contests in 1955.

No other type of contest occurred more than half a dozen times and, between them, these three categories accounted for 530 of the 630 seats. At first all constituencies were treated alike, but as soon as sufficient results were available, separate swings were estimated for each group; these were then applied to predict other seats in the same group. Seats outside these three groups were not used to estimate swings, but their results were predicted by applying to the three main parties the swings obtained from the most similar of the three types of contest.

Thus, at 2.45 a.m. on Friday morning the estimated swings in group 2 were: Conservative −1·1%, Labour −3·1%, Liberal +4·2%. The 1955 results in, for example, Cornwall North, were:

Conservative 47·4%, Labour 9·8%, Liberal 42·8%

Adding the estimated swings to the 1955 results gives:

Conservative 46·3%, Labour 6·7%, Liberal 47·0%.

In fact this was one of the two seats in which the program (and the Liberals!) wrongly predicted a Liberal victory; the actual results were:

Conservative 46·6%, Labour 9·5%, Liberal 43·9%.

An analysis of the last election (Butler, 1955) had shown that in 1955 simple party swings were insufficient for accurate forecasting due to the presence of a swing towards the sitting member. The program therefore estimated this swing from the straight fight results, and then applied it when predicting any seat in which the previous M.P. was standing.

At 2.45 a.m. on Friday the overall swing to Conservative in straight fights was +1·3%. Where the previous Socialist M.P. stood the average swing was +1·0%; where the previous Conservative M.P. stood it was +1·6%. Thus the swing to the sitting member was estimated to be +0·3%, the same figure as Butler found in 1955. If the Conservative candidate had been the previous M.P. in Cornwall North, this swing would have been added to his estimated votes and subtracted from those of his nearest rival, giving the prediction:

Conservative 46·6%, Labour 6·7%, Liberal 46·7%

Several other factors were calculated in the analysis of the results, but it was thought unwise to use them for prediction. Separate swings were calculated for nine different regions, but there were too few results in each category to regard these as reliable figures for prediction. For the same reason marginal seats were not separately predicted. A few seats with large boundary changes, such as Plymouth, or with unusual contests in either election, such as Gravesend, were not used for calculating the swing. When making a prediction the computer took the results already known and added to them its own predictions for the remaining seats. It was never supposed that any program could predict all seats correctly, but in the event the individual errors did cancel out, as expected, to give very good forecasts of party strengths.

THE RESULTS

After the first approximation, known to be based on too few results, the program settled down to give a forecast which proved to be consistently close to the final state of the parties (see Table 1). Soon after 11.0 p.m., when 17 results were available, the computer predicted a Conservative representation of 359. This rose to 364 by 11.50 p.m. and fluctuated only slightly to reach 368 by 2.45 a.m., when the last prediction was made. This figure differed by three from the final figure of 365; one of these was Sir David Robertson, the Independent in Caithness, who had been included as a Conservative.

The small differences between the prediction and the final result can be considered to justify the original assumption that abnormal results would tend to cancel out.

As well as making predictions the computer was used to analyse the results by regions and large towns. This showed the percentage poll, the party strengths, gains and losses, percentage votes, and swings for each party. It was apparent from these that there were regional differences in swing: Scotland and the north-west swung towards Labour, whilst the London area had a stronger swing towards Conservative than the rest of the country. Some of these regional swings fluctuated rather wildly at first and would have been useless for prediction.

TABLE 1

SOME OF THE PREDICTIONS

Time	*Results available*	*Prediction*			
		Conservative	*Labour*	*Liberal*	*Other*
10.40 p.m.	6	309	313	6	2
10.45 p.m.	8	342	280	6	2
11.00 p.m.	17	359	265	6	0
11.50 p.m.	63	364	260	6	0
12.30 a.m.	150	366	257	7	0
1.30 a.m.	291	367	255	8	0
2.45 a.m.	363	368	254	8	0
RESULT	630	365*	258	6	1*

* Note that the Independent, who had been a Conservative, was included in the prediction as a Conservative.

CONCLUSIONS

It has been common practice to apply computers only to the routine calculations of swings. The results indicate that a computer can be used directly to make accurate forecasts, thus becoming something more than just a calculating machine at the elbow of the political expert.

More interesting is the fact that a satisfactory method was devised in this case for rapidly assimilating and checking relatively unreliable data. Data checking is of vital importance in computer applications, and it may be that similar methods could be usefully applied in other fields.

ACKNOWLEDGEMENTS

The authors wish to express their thanks to their many colleagues who assisted both before and during the election. This paper is published by permission of Messrs. Ferranti Ltd.

REFERENCES

BUTLER, D. (1955). *The British General Election of* 1955. London, Macmillan and Co. Ltd.; New York, St. Martin's Press.

———— (1959). "Unprecedented Accuracy in Predicting Results," article by Scientific Correspondent, *The Guardian*, Manchester, 10 October 1959.

LEO Automatic Office at the Ministry of Pensions and National Insurance

The report is now available on the commissioning tests carried out recently on the LEO Automatic Office installed in the offices of the Ministry of Pensions and National Insurance in Newcastle.

The tests were set by an expert body formed to advise Government departments on technical standards of computers that they might be considering.

The test period was the whole week of 2–6 November, starting at 9.30 and continuing until 17.30, save on the last day when work stopped at 17.00. A series of 9 tests was run repetitively, 5 of them diagnostic test programs designed to examine exhaustively the correct functioning of all aspects of the equipment, and 4 of them actual programs taken from the payroll suite that is planned to be the first operational work taken over by the installation.

The trials were under control of the Ministry personnel, trained by LEO to operate the equipment. Careful note was taken of any excess time on these so as to dissect it between that which was attributable to a fault in any part of the computer assemblage and that which was due to operational inefficiency of any sort. Operational inefficiency included data-preparation faults, such as mis-sorts of cards, feeding of stationery, and any odd minutes lost through operator mishandling.

The final outcome of the week was as follows:—

		Hours	*Minutes*
A.	Good running time	37	44
B.	Time lost due to technical causes (including re-runs)		51
C.	Time lost due to operating inefficiencies		54
		39	29

The criterion set was based on $A/(A - B)$. On this basis LEO achieved over the whole exacting week a result of 97·8%, which very satisfactorily exceeded the standards looked for.

The equipment includes a LEO II Computer with a 16,000-word magnetic drum in addition to the 2,048-word high-speed store, three concurrent punched-card input channels with provision for a fourth, two punched-card output channels, and a printer output channel.

Following the normal LEO arrangements, the programming of the first suite of jobs has been proceeding side by side with the building of the equipment. Ministry programmers, selected in all cases from existing Ministry staff, and trained by LEO, form the greater part of the team at work on the suite, but the responsibility for the successful completion of the suite has been undertaken by LEO Computers Limited. The work has been supervised by a Senior Programmer from LEO, who has combined the technical progressing of the jobs with the further practical training of the Ministry programmers.

It is planned that the first jobs will commence operational running in January and the load will steadily build up until all the 3,000 weekly paid officers stationed at Newcastle, and the 20,000 monthly paid officers at local and regional offices scattered throughout England, will be paid through the LEO installation. Other work planned includes the production of analysed health statistics, derived from accident and sickness records, and the control of issue and receipt of National Insurance benefit claims.