AJPH
AUSTRALIAN JOURNAL OF POLITICS AND HISTORY

Australian Journal of Politics and History: 2024.

Using Numerical Anomalies to Test for Fraud in Colonial New South Wales Elections

ANDREW LEIGH
Parliament of Australia

When fabricating numbers, humans tend to make systematic errors, favouring some numerals over others. Analysing twenty-first-century elections in which fraud has been alleged, researchers have observed anomalous patterns in the last digits of vote counts. I show that vote counts for the 2022 Australian election exhibit no anomalies, and then turn attention to nineteenth-century colonial elections in New South Wales. Pooling data for elections from 1843 to 1887, I find a less uniform pattern in final digits. Even so, formal statistical tests cannot reject the hypothesis that vote counts in these elections were unaffected by fraud.

In 2012, New York University's Bernd Beber and Alexandra Scacco published a seminal paper that analysed the numbers in electoral results for various countries.[1] Their starting point was that humans are surprisingly bad at making up numbers. When asked to create random numbers, participants have been shown to favour some numerals over others. Numbers ending in zero do not feel random to us, so less than one-tenth of made-up numbers end in zero. Falsified numbers are more likely to end in one and two than in eight and nine.

Using this insight, the researchers look at the last digit of vote counts in various elections — both in Scandinavia and in Africa. They find that vote counts in the 2002 Swedish election have last numbers that are evenly distributed from zero to nine. By contrast, in elections held in Nigeria in 2003 and Senegal in 2007, they find anomalous patterns of last digits. Other researchers have used the tools of forensic electoral analysis, applying them to modern-day elections where fraud has been alleged.[2]

In this short research note, I apply the same approach to Australia. I begin by studying the 2022 election, conducted by the highly regarded Australian Electoral

These results were first presented to the Australian Government Data Summit in Canberra on 28 March 2024. My thanks to the Parliamentary Library's Christopher Guiliano and my staff members Bria Larkspur, Angella Fernando, Georgia Thompson, and Maria Neill for assistance with data-crunching, and to Antony Green for his original tabulations of NSW colonial election results. Frank Bongiorno provided valuable feedback on an earlier draft. The author has no conflicts of interest [Correction added on 15 May 2024, after first online publication: The acknowledgements section has been updated to reflect the correct name of a staff member.].

[1] Beber, B. and Scacco, A., "What the Numbers Say: A Digit-Based Test for Election Fraud," *Political Analysis*, Vol 20, 2 (2012), pp. 211–34.

[2] J. Deckert, M. Myagkov and P. C. Ordeshook, "Benford's Law and the Detection of Election Fraud," *Political Analysis*, Vol 19, 3 (2011), pp. 245–68; D. Figueiredo Filho, L. Silva, and E. Carvalho, "The Forensics of Fraud: Evidence from the 2018 Brazilian Presidential Election," *Forensic Science International: Synergy*, Vol 5 (2022), p. 100286; L. Pericchi and D. Torres, "Quick Anomaly Detection by the Newcomb—Benford Law, with Applications to Electoral Processes Data from the USA, Puerto Rico and Venezuela," *Statistical Science*, Vol 26, 4 (2011), pp. 502–16; S. Tunmibi and W. Olatokun, "Application of Digits Based Test to Analyse Presidential Election Data in Nigeria," *Commonwealth and Comparative Politics*, Vol 59, 1 (2021), pp. 1–24.

DOI:10.1111/ajph.12988.

Commission, and characterised by a range of integrity and transparency mechanisms, including scrutineers, measures to protect voter safety, a mandatory second count, security protocols around the handling of ballot papers, and oversight from the Electoral Integrity Assurance Taskforce.

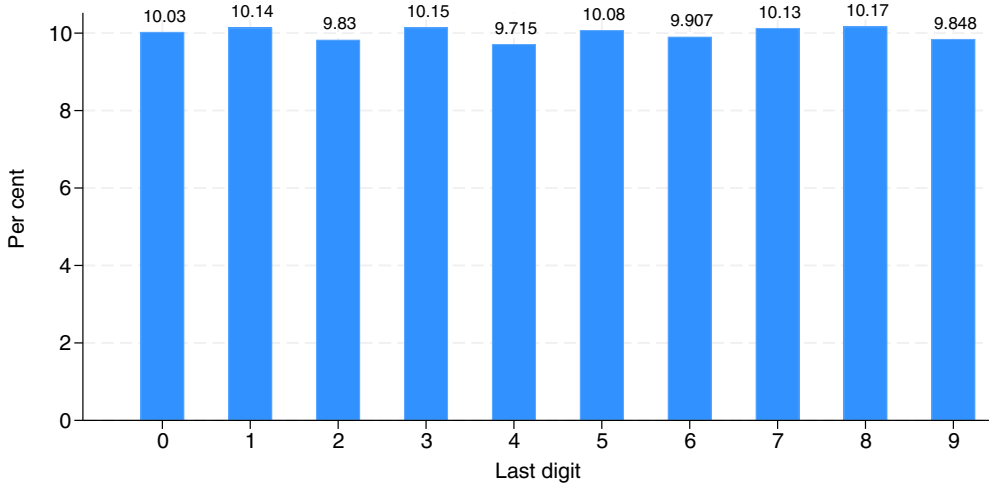
Having established that the last digits of vote counts in the 2022 Australian election were uniformly distributed, I study the results of elections held in Australia from 1843 to 1887. To preview the results, my analysis of the last digits of these vote counts finds that they are less uniform than the 2022 Australian election. However, a formal statistical test does not allow me to reject the hypothesis that the last digits are uniformly distributed.

Studying the Distribution of Last Digits in the 2022 Australian Election

As a falsification test of the methodology, I first study the 2022 Australian Election. No serious allegations of fraud have been made regarding this election. Two months after the election, the Australian Electoral Commission's Electoral Integrity Assurance Taskforce reported that it had identified no interference that compromised the delivery of the 2022 federal election.[3]

To analyse last digit patterns in the 2022 election, I use all candidate vote counts at a polling place level. When vote counts are low, the last digit tends to be smaller. Accordingly, I omit results where a candidate received one hundred or fewer votes. That leaves 21,822 results. Figure 1 shows the distribution of last digits across these vote counts. The last digits show a refreshingly uniform pattern. The least common final digit is four, which appears 9.7 per cent of the time, and the most common final digit is eight, which appears 10.2 per cent of the time. A chi-square test fails to reject

Figure 1 Last Digits in Vote Counts for 2022 Australian Election (Where Vote Counts Exceed 100)
[Colour figure can be viewed at wileyonlinelibrary.com]



[3] Australian Electoral Commission, "Taskforce Advice Provides Assurance on 2022 Federal Election," Media Release, 26 July 2022.

the hypothesis that the distribution of last digits is uniform (the test statistic is 5.2, with a p-value of 0.8).

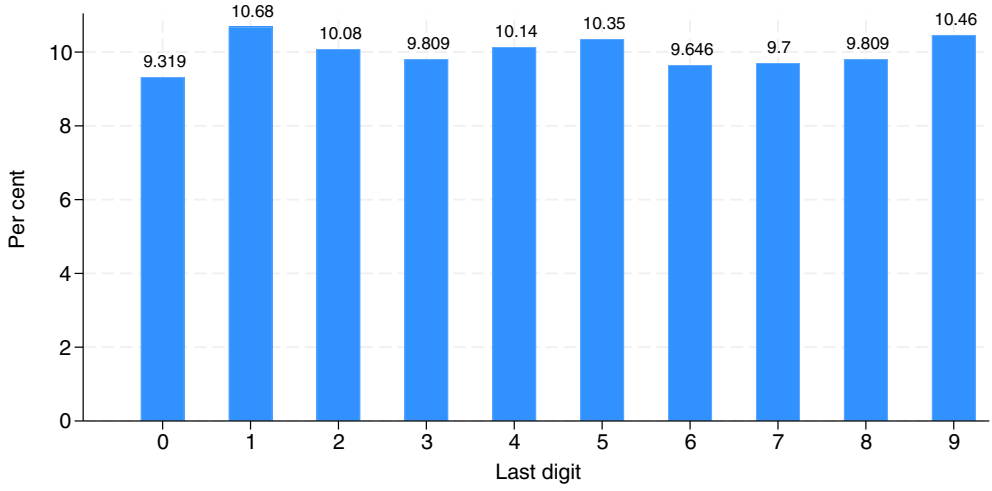
Studying the Distribution of Last Digits in New South Wales Elections, 1843–87

The first colony-wide election in Australia took place for the New South Wales (NSW) Legislative Council in 1843. Only men who owned land worth more than £200 or more (or who rented a dwelling for £20 or more a year) were permitted to vote. According to the National Museum of Australia, these early Australian elections were "rough and tumble exercises […] Alcohol, bribery, coercion and violence were intrinsic to the process".[4] Men were shot, and the Irish journalist William Kelly described early colonial polls as "nothing more or less than pantomime in a frenzy". Prior to 1856, the NSW parliament was unicameral, with the Legislative Council being partially elected. With the creation of the fully-elected Legislative Assembly in 1856, the Legislative Council became a fully appointed body.

What can be observed about the voting patterns in these early elections? Because there were relatively few contests, I begin by analysing them together, combining data for the first sixteen NSW colonial elections, from 1843 to 1887. There is no obvious cut-off point for "early" elections, but I choose 1887 on the basis that it was the last election before payment for members of parliament was introduced. I include both Legislative Council and Legislative Assembly elections, as well as by-elections. Dropping contests in which candidates received 100 votes or fewer leaves 1,835 results. The distribution of the last digits is shown in Figure 2.

With fewer races in the dataset, it is reasonable to expect more statistical noise. Nonetheless, the digit frequencies all lie in the range from 9 to 11 per cent. The least common last digit is zero, which appears 9.3 per cent of the time, while the most

Figure 2 Last Digits in Vote Counts for 1843–87 NSW Elections (Where Vote Counts Exceed 100)
[Colour figure can be viewed at wileyonlinelibrary.com]



[4] National Museum of Australia, "Secret Ballot Introduced," 2024, *https://www.nma.gov.au/defining-moments/resources/secret-ballot-introduced*. See also J. Brett, *From Secret Ballot to Democracy Sausage: How Australia Got Compulsory Voting* (Text Publishing, Melbourne, 2019), pp. 12–14.

common is one, which appears 10.7 per cent of the time. The underrepresentation of numbers ending in zero and the overrepresentation of numbers ending in one is consistent with a tendency of humans to fabricate numbers in this fashion. However, the difference is small, and a chi-square test fails to reject the hypothesis that the overall distribution of last digits is uniform (the test statistic is 2.9, with a p-value of 0.9).

Finally, I carry out separate chi-square tests for each individual election. These fail to reject the hypothesis of uniformity in all instances. The 1856 and 1859 elections come closest to deviating from uniformity, with p-values on the chi-square test of 0.11 and 0.12, respectively.

Conclusion

Detecting election fraud from unexpected patterns in polling results is an approach that has borne fruit when analysing electoral results from potentially fraudulent twenty-first-century elections. However, when applied to nineteenth-century Australian elections, anecdotal allegations of electoral irregularity are not reflected in highly unusual patterns of last digits.

However, it would be a mistake to conclude from this that there were no electoral anomalies in colonial NSW elections. The last digit approach to detecting fraud was developed to uncover electoral irregularities in national presidential elections, where a candidate's supporters are able to take control of polling places and directly manipulate vote totals. The last digit approach would not pick up other kinds of electoral misconduct. For example, if some voters are intimidated into not attending a polling place, but the counts are nonetheless carried out by an independent body, then the last digits would not be expected to show any deviation from uniformity.

Likewise, in the case of NSW colonial elections, it may be the case that misconduct was chaotic rather than organised. In this instance, the election results may not have fully reflected the preferences of the electorate, yet the last digits of vote counts could be reasonably uniform.

Nonetheless, the last digit approach may prove useful in other historical contexts. If historians identify specific elections — in Australia or elsewhere — that may have been marred by fraud, then it might be possible to test for numerical anomalies. This could involve unusual patterns in last digits, as analysed in this paper, or other numerical anomalies, such as too few repeated numbers or a pattern of first digits that does not conform to Benford's Law.[5] Just as the analysis of unexpected numerical patterns has been used in cyber forensics, forensic economics, and forensic accounting, these techniques may also be applicable to the study of irregularities in historical elections.

[5] Benford's Law suggests that the distribution of first digits will follow the rule $P(d) = \log_{10}(1 + 1/d)$, where d ranges from 1 to 9. For example, $P(1) = 0.3010$ and $P(9) = 0.0458$.