B. DISCUSSION OF PREVIOUS RESEARCH

In the late 1990s and early 2000s, the Office of Research Integrity (ORI, a division of the U.S. Department of Health and Human Services) used a set of tools similar to the one we propose, and ORI-affiliated researcher James E. Mosimann and a number of co-authors published three articles that provide foundations for and applications of this approach (Mosimann, Wiseman, and Edelman, 1995; Mosimann and Ratnaparkhi, 1996; Mosimann et al., 2002). Two of the articles appeared in the specialty journal *Accountability in Research*, which focuses on research in medical ethics. The third article appeared in *Communications in Statistics – Simulation and Computation*. None of the articles have been widely cited and they have to our knowledge not been cited by any political scientists.[1]

We independently developed a set of fraud detection tools similar to those developed by Mosimann et al., and we make at least three contributions that go beyond their work.

First, we provide alternative theoretical foundations for the last-digit test, which reflect more directly the nature of electoral returns. Mosimann and Ratnaparkhi (1996) prove the theoretical result that the rightmost digits in the decimal expansion of a continuous random variable will be approximately uniformly distributed, while we prove that the last digits of integer realizations of a discrete random variable are uniformly distributed under certain conditions. ORI investigations frequently involve numbers drawn from continuous distributions such as coefficients or recorded weights, while we focus on integer vote counts.[2]

We also expand on Mosimann and Ratnaparkhi (1996) in that we explicitly prove our result for any positional numeral system, while their analysis focuses on digits in base-10.

[1] According to Google Scholar and excluding references among the three articles themselves, Mosimann, Wiseman, and Edelman (1995) has been cited eight times, exclusively in relation to investigations of scientific misconduct, in journals such as *Mutation Research*, *Cancer Risk Evaluation*, and *Pharmaceutical Statistics*. Mosimann and Ratnaparkhi (1996) has been cited twice, once in *Science and Engineering Ethics* and once in the *Journal of Quantitative Criminology*, and Mosimann et al. (2002) has been cited four times. In the course of the review process, we became aware of Mosimann et al.'s work through a reference in Diekmann (2007).

[2] Although the proofs in Mosimann and Ratnaparkhi (1996) and our paper proceed differently, there are similarities in intuition; see in particular Mosimann and Ratnaparkhi (1996, 504–5).

Although vote returns are naturally written in decimal notation, this means that our result extends for example to numbers in binary notation.

Second, we move beyond Mosimann et al.'s work in that we introduce tests that examine pairs of trailing digits to complement our last-digit analysis. We derive new theoretical and empirical expectations with respect to the distance between last and penultimate digits and suggest the application of this approach alongside the last-digit test.

Third, we open up a new field of application for the last-digit test by presenting it in the context of political science and the study of elections. We address issues and challenges faced by a digit-based approach that are particular to this field of study, such as the effect of aggregation across different levels of tabulation, or the challenge of distinguishing rounded vote counts from fraudulently manipulated results.

REFERENCES

Diekmann, Andreas. 2007. "Not the First Digit! Using Benford's Law to Detect Fraudulent Scientific Data." *Journal of Applied Statistics* 34(3):321–329.

Mosimann, James E., and Makarand V. Ratnaparkhi. 1996. "Uniform Occurrence of Digits for Folded and Mixture Distributions on Finite Intervals." *Communications in Statistics - Simulation and Computation* 25(2):481–506.

Mosimann, James E., Claire V. Wiseman, and Ruth E. Edelman. 1995. "Data Fabrication: Can People Generate Random Digits?" *Accountability in Research* 4(1):31–55.

Mosimann, James E., John E. Dahlberg, Nancy M. Davidian, and John W. Krueger. 2002. "Terminal Digits and the Examination of Questioned Data." *Accountability in Research* 9(2):75–92.