Newcomb–Benford law helps customs officers to detect fraud in international trade

Lucas Lacasa[a,1]

The leading digit of a number represents its nonzero leftmost digit. For example, the leading digits of 19 and 0.072 are 1 and 7, respectively. The Newcomb–Benford law (NBL) was originally discovered in the late 19th century (1, 2) as an anecdotal pattern emerging in such seemingly disparate datasets as streets addresses, freezing points of chemical compounds, house prices, and physical constants, with the leading digit, d, in those datasets following a logarithmically decaying distribution, $P(d) = \log_{10}(1 + 1/d)$, instead of being uniformly distributed, as one may naively assume. Later, this pattern was shown to be a consequence of a central limit-type mechanism (3–5), emerging not only empirically but also in mathematical sequences of several garments. A few years ago, some authors devised a way to leverage the NBL as an antifraud tool (6, 7), based on a simple idea: Assuming that this law is expected to naturally emerge in a certain dataset, the statistics would deviate from the law in a way that could be quantitatively measured when the dataset has been manipulated or when data have been fabricated. Accordingly, the NBL and variants have been proposed to assess fraud in contexts ranging from election data (8–11) to financial accounting in external, internal, and governmental auditing (12). In PNAS, Cerioli et al. (13) take this strategy to the next level, proposing a sophisticated statistical modeling framework that can be used to monitor and detect hints of individual fraudulent behavior in the context of international trade (i.e., imports and exports that are declared by national traders and shipping agents). Cerioli et al. developed a mathematical model that provides the correct statistical tests to assess conformance of individual traders to the NBL and then validated the whole framework in realistic scenarios enabled by high-resolution transaction data from the customs of various European Union (EU) member states.

A naive approach to flagging potential misuse or manipulation of data using the NBL is the following: under the premise that a particular dataset conforms to a specific theoretical distribution (NBL or other variants of the law; e.g., the second leading digit), the strategy is to compare the empirical leading-digit distribution found in the actual data against the expected NBL. Applying goodness of fit and a contrast of hypothesis, one can conclude whether the null hypothesis (i.e., the empirical one conforms to the expected NBL) can be rejected up to a certain confidence level. This is typically the approach taken to assess vote count statistics in the context of election forensics (8–11). Note, however, that methodologies based on the NBL and variants are not totally free of healthy controversy (14, 15) and that potential fraudsters aware of these regularities can try to make up data in such a way that specific patterns, which emerge in fair data, also hold in the manipulated data. Even more dramatically, despite some common misperception that the origin of the NBL is reasonably well understood, this is not actually the case (16). In other words, there is no general theory fully guaranteeing that the NBL should systematically emerge in a particular unmanipulated dataset.

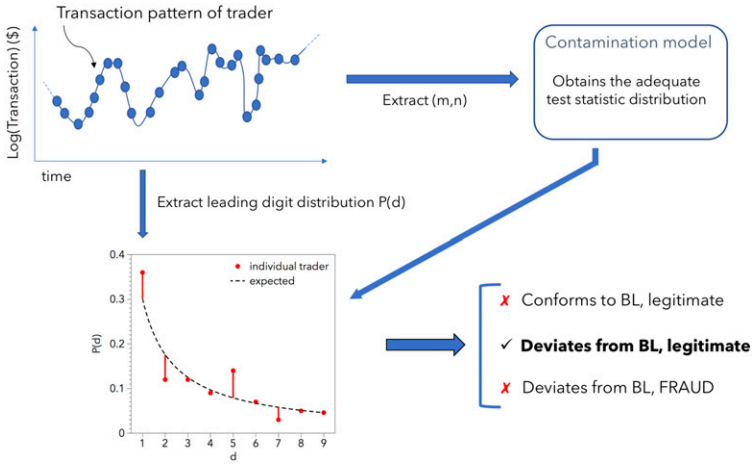

Fig. 1. An individual trader leaves a trace of transaction activities. The method computes the leading digit distribution, $P(d)$, and (*i*) evaluates which is the expected law if no manipulation has occurred and (*ii*) obtains the adequate test statistic distribution. Statistical comparison between the empirical and expected distribution concludes whether the data conform or deviate from the NBL, and, in the latter case, whether or not this deviation is due to legitimate reasons. BL, Benson's law.

[a]School of Mathematical Sciences, Queen Mary University of London, E14NS London, United Kingdom

Author contributions: L.L. wrote the paper.
The author declares no conflict of interest.

See companion article on page 106.

[1]Email: l.lacasa@qmul.ac.uk.

Published online December 12, 2018.

In their report, Cerioli et al. (13) are able to circumvent these issues by preassessing the conditions under which the NBL should emerge in the context of international trade—that is, the conditions under which the subsequent inferences can be trusted. These authors actually pioneer the application of the NBL to the context of international trade. Why is international trade an important area of application? Because of the money. Big money. To give a sense of the scale, the EU accounts for around 15% of the world's trade in goods, translating to over €1,800 billion on imports and a similar quantity of exports in 2017 alone, with a positive balance between the two of over €20 billion (17). Being able to detect fraudulent behavior—including underreported goods—is therefore of paramount importance (18). Devising a robust and accurate statistical method that could be easily embedded online to automatically monitor transaction activities and flag suspicious ones would tick all of the boxes (i.e., reliability, flexibility, efficiency, and cost-effectiveness). This approach also aligns with the overall modernization strategy currently pursued by the World Customs Organization (19).

Cerioli et al. (13) successfully address the two main challenges that might otherwise preclude a realistic detection of fraud via NBL analysis of individual traders. First, the authors establish solid conditions for the validity of the NBL in the field of international trade data—an essential step for the implementation of large-scale, automatic monitoring processes. Second, they find approximations and corrections to the adequate test statistics needed to scrutinize fraud in those instances in which the NBL is not actually expected to hold. Importantly, the authors are therefore able to discriminate two types of nonconformance to the NBL: cases related to data fabrication and fraud versus the so-called false-positives, which are legitimate deviations that emerge, for example, for traders who operate on a limited number of products (so that there is not enough variability in their transactions for the NBL to emerge in the first place). As a result, traders can be classified into three groups: (*i*) legitimate traders whose activity conforms to the NBL, (*ii*) traders whose activity is still legitimate but does not conform to the NBL due to controlled factors, and (*iii*) traders whose activity does not conform to the NBL even when it should, probably due to data fabrication (see Fig.1 for an illustration).

On a technical level, Cerioli et al. (13) initially argue that the sequence of transaction values for a particular trader complies with Hill's (3) central limit theorem hypothesis. By generating synthetic transactions of "idealized" traders, the authors can statistically assess the conditions where the NBL holds and in what circumstances classic χ^2 tests can therefore be applied. An important result is that the NBL breaks down as the theoretical expected distribution when the number of different traded goods, m, is much smaller than the total number of transactions, n ($m \ll n$), so that in those instances, a blind χ^2 test is not recommended.

One of the important aspects of Cerioli et al.'s report (13) is that they model transaction data via a trader-specific contamination model composed by linearly interpolating a mixture of two distributions for the leading digit: a legitimate one (usually the NBL), which is parametrically dependent on n and m; and a contaminant distribution, which models the effect of data manipulation. The null hypothesis (legitimate trader) is equivalent to having a null coefficient on the contaminant distribution; however, this scenario is equivalent to conformance to the NBL only when the transaction data fulfill the criteria discussed above. In those cases in which the expected distribution for nonfraudsters deviates from the NBL for legitimate reasons, Cerioli et al. can plug in the actual expected distribution as the legitimate one and can deduce, among other things, the correct distribution of the test statistic that shall be used to reject the null hypothesis of a legitimate transaction pattern.

Cerioli et al.'s paper provides a principled framework for goodness-of-fit testing of the NBL for antifraud purposes, with a focus on customs data.

Validation and calibration of the theoretical framework was possible thanks to access to real data of import/exports declared by national traders and shipping agents using the Single Administrative Document form (data were provided by the Italian Customs Agency and by the customs office of another European Union member state not disclosed for its specific confidentiality policy). Results overall show that Cerioli et al.'s (13) methodology can flag fraudsters as well as deflag traders whose activity deviates from expected the NBL due to legitimate reasons. The authors discuss a particularly illuminating case of a trader extracted from an archive of fraudulent declarations provided by the Italian Customs Agency, whose fraudulent behavior was discovered only after substantial investigation of two of the declarations. In this case, a standard protocol based on robust regression techniques aiming at the automatic detection of value frauds in customs data did not provide clear evidence of substantial undervaluation or of other major anomalies. Cerioli et al.'s analysis, on the other hand, produced a strong signal of contamination of the digit distribution for this trader, and their statistical analysis safely concluded the presence of fraudulent manipulation.

In conclusion, Cerioli et al.'s paper (13) provides a principled framework for goodness-of-fit testing of the NBL for antifraud purposes, with a focus on customs data. This methodology has the potential to be embedded—probably in combination with more standard and model-free approaches (6, 12)—in real international trade antifraud protocols and audits in the near future. In this respect, a web application (13) developed with the purpose of assisting customs officers and auditors in the screening task has already been set in place.

Acknowledgments

L.L.'s research is supported by Engineering Physical Sciences Research Council Early Career Fellowship EP/P01660X/1.

1 Newcomb S (1881) Note on the frequency of use of the different digits in natural numbers. *Am J Math* 4:39–40.
2 Benford F (1938) The law of anomalous numbers. *Proc Am Philos Soc* 78:551–572.
3 Hill TP (1995) A statistical derivation of the significant-digit law. *Stat Sci* 10:354–363.
4 Berger A, Hill TP (2015) *An Introduction to Benford's Law* (Princeton Univ Press, Princeton).
5 Miller SJ, ed (2015) *Benford's Law: Theory and Applications* (Princeton Univ Press, Princeton).
6 Kossovsky AE (2015) *Benford's Law: Theory, The General Law Of Relative Quantities, and Forensic Fraud Detection Applications* (World Scientific, Singapore).
7 Lacasa L, Fernández-Gracia J (2018) Election forensics: Quantitative methods for electoral fraud detection. *Forensic Sci Intl*, 10.1016/j.forsciint.2018.11.010.
8 Mebane WR (2006) *Election Forensics: Vote Counts and Benford's Law* (Political Methodology Society, Univ California, Davis, CA).

9 Pericchi L, Torres D (2011) Quick anomaly detection by the Newcomb-Benford Law, with applications to electoral processes data from the USA, Puerto Rico, and Venezuela. *Stat Sci* 26:502–516.
10 Fernández-Gracia J, Lacasa L (2018) Bipartisanship breakdown, functional networks, and forensic analysis in Spanish 2015 and 2016 national elections. *Complexity* 2018:9684749.
11 Breunig C, Goerres A (2011) Searching for electoral irregularities in an established democracy: Applying Benford's law tests to Bundestag elections in Unified Germany. *Elect Stud* 30:534–545.
12 Nigrini MJ (2012) *Benford's Law* (Wiley, Hoboken, NJ).
13 Cerioli A, Barabesi L, Cerasa A, Menegatti M, Perrotta D (2019) Newcomb–Benford law and the detection of frauds in international trade. *Proc Natl Acad Sci USA* 116:106–115.
14 Deckert JD, Myagkov M, Ordeshook PC (2011) Benford's Law and the detection of election fraud. *Polit Anal* 19:245–268.
15 Mebane WR (2011) Comment on "Benford's Law and the detection of election fraud". *Polit Anal* 19:269–272.
16 Berger A, Hill TP (2011) Benford's law strikes back: No simple explanation in sight for mathematical gem. *Math Intell* 33:85–91.
17 Eurostat Statistics Explained (2017) International trade in goods—A statistical picture. Available at https://ec.europa.eu/eurostat/statistics-explained/index.php?title=International_trade_in_goods_-_a_statistical_picture. Accessed November 23, 2018.
18 European Commission (2014) Operation SNAKE: EU and Chinese customs join forces to target undervaluation of goods at customs. Available at europa.eu/rapid/press-release_IP-14-1001_en.htm. Accessed November 23, 2018.
19 World Customs Organization (2017) Message from the WCO Secretary General. Available at www.wcoomd.org/en/media/newsroom/2017/january/message-of-the-wco-secretary-general-on-international-customs-day-2017.aspx. Accessed November 23, 2018.