Application of Benford’s law in Data Analysis

Feifei Li, Shuqing Han, Hongyv Zhang, Jiaojiao Ding, Jianhua Zhang and Jianzhai Wu*

Agricultural Information Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, China

*Corresponding author: wujianzhai@caas.cn

Abstract. In the era of big data, data is growing explosively. Data quality control has become a key factor in maximizing data value. It is important and urgent to establish a scientific data quality detection method. Benford’s law has become an effective tool for detection of data quality and identification of anomaly data in various fields. On the basis of expounding the basic principles of Benford’s law, this paper summarizes its application in different levels of natural and social sciences, explores the data conditions of the law and the factors affecting the accuracy of the detection, and proposes the model improvement idea in three aspects of extending detection scope, combination of several methods and strengthening the explanation of detection results. The development of Benford’s Law in the future requires scholars from all fields to study more about its essence, strengthen its integration with other data processing technologies, and then expand its application.

1. Introduction

In recent years, the convergence of information technology with several fields of economic society has stimulated the explosive growth of data volume. According to IDC’s report, the world entered ZB era in 2010 and is currently growing at a rate approximately doubling every two years. It is expected that the global data will reach 35ZB by 2020. The era of big data has arrived[1,2].

While the data volume is growing rapidly, data quality problems such as junk data, erroneous data, and false data are increasingly obvious[3], which affect the knowledge discovery and value innovation. Big data must not only have “big volume” but also have “high quality”. In the “Outline for the Promotion of Big Data Development” issued by the State Council in 2015, the simultaneous improvement of data scale, quality and application level is required to be achieved, so the exploration of efficient and reliable data quality is very important. The first digit described by Benford’s law satisfies a uniform logarithmic distribution, and compares it with the actual distribution of data and analyze, the purpose of testing data quality and mining anomaly data can be achieved. It can also be combined with various theoretical techniques to solve other practical application problems. It has been widely used in the fields of natural and social sciences.

2. Development of Benford’s Law

2.1 Generation and development

In 1881, American astronomer and mathematician, Simon Newcomb, first discovered the law of the first digit. Through statistical analysis of the data, he pointed out that the numbers in nature can be expressed in exponential form, and the logarithm of the index mantissa is uniformly distributed[4]. But
at the time his research did not attract the attention of the academia. Until 1938, American physicist, Frank Benford, once again discovered the same phenomenon, and conducted deep research on 20229 digits in more than 20 data sets appearing randomly in physical and chemical constants, prime numbers and Fibonacci numbers, river length and lake area, census data, and even books or magazines, and got the same conclusions as Newcomb, and proposed the "First Digit Law " formally[5]. People from the academia began to have a strong interest in it. Later this law was named as "Benford’s Law."

In the next 40 years, the development of basic research was very slow. Although the related research explained some of the properties of Benford’s law from a mathematical point of view, its mathematical principle was not proved.

Until 1995, American scholar, Hill, used the form of the effective numerical center limit theorem in statistics creatively. For the first time, he theoretically proved that the first digit obeyed the principle of Benford’s law[6,7]. In addition, Hill also generalized the theory and obtained the distribution function law of high order digit, and derived the joint distribution function between the first digit and the higher order digit. So far, the basic theories and methods of Benford’s law have been initially improved.

2.2 Form

Benford’s law, also known as the "First Digit Law", means that in nature, the probability of the first digit being 1-9 is unequal, and it satisfies a uniform logarithmic distribution. The distribution probability of the first digit being 1-9 is shown in Equation 1:

$$P(D = d_1) = \log(1 + \frac{1}{d_1}) \quad (d_1 = 1, 2, \ldots, 9) \quad \text{equation (1)}$$

| First digit | 1      | 2      | 3      |
|-------------|--------|--------|--------|
| Probability | 0.30103| 0.17609| 0.12494|
| First digit | 4      | 5      | 6      |
| Probability | 0.09691| 0.07918| 0.06695|
| First digit | 7      | 8      | 9      |
| Probability | 0.05799| 0.05119| 0.04576|

$d_1$ is the first non-zero valid integer in the number. For example, the first digit of 245 and 0.004 is 2, 4, respectively. $P(D = d_1)$ refers to the probability that the first digit is $d_1$, $P(d_1=1)=0.301$, $P(d_1=2)=0.176$, and so on, the first digit probability distribution table is shown in Table 1. Obviously, the probability of the first digit being 1 to 9 is not simple $1/9$, but close to 3 times the expected value of $1/9$; the whole shows the law of monotonous decline; the larger the digit is, the probability of its appearing as the first digit is lower.

The second and third digit distribution probability is shown in Equation 2 and Equation 3. The calculation shows that the probability that the second and third digit is 0 to 9 is also in descending order, and the corresponding decline of the digit is much smaller than the distribution law of the first digit, and it gradually slows down. The probability distribution comparison of figure 1-9 in the top three digits is shown in Figure 1.

$$P(D_2 = d_2) = \sum_{d_1=1}^{9} \log(1 + \frac{1}{d_1d_2}) \quad (d_2 = 0, 1, 2, \ldots, 9) \quad \text{equation (2)}$$

$$P(D_3 = d_3) = \sum_{d_1=1}^{9} \sum_{d_2=1}^{9} \log(1 + \frac{1}{d_1d_2d_3}) \quad (d_3 = 0, 1, 2, \ldots, 9) \quad \text{equation (3)}$$
Fig.1 Comparison of digital frequency distribution of figure1-9 in the top three digits

The joint distribution function probability of the first nth digit is shown in Equation 4. For example, the joint occurrence probability of 218 in the first, second and third digit is $\log_{10}(1 + (218)^{-1})$. According to the joint distribution function and the multiplier rule, we may infer that the probability distribution function of the jth digit is $d_j$ when the known ith digit is $d_i$, as shown in Equation 5.

\[
P(D_1 = d_1, D_2 = d_2 \ldots D_n = d_n) = \log_{10}(1 + \frac{1}{d_1 d_2 \ldots d_n}) \quad \text{equation (4)}
\]

\[
P(D_1 = d_1, D_2 = d_2 \ldots D_n = d_n) = \frac{P(D_2 = d_2, D_3 = d_3) \ldots P(D_n = d_n)}{P(D_1 = d_1)} = \log_{10}(1 + \frac{1}{d_1})
\]

\[
\text{equation (5)}
\]

3. Application of Benford’s law

Benford’s law is an efficient digital analysis technology. The application in data analysis can be roughly divided into three levels: applicability verification, anomaly data detection, and cross-application with other methods.

3.1. Applicability verification

Applicability verification means verification of the usability of Benford’s law in different fields. Currently, relevant research has been carried out in the fields of economics, sociology, physics, computer science and biology. Among them, the most influential research is the systematic applicability analysis of tax and accounting and other data by Nigrini in the 1990s, which not only clarified the availability of the law in the relevant financial directions [8], but also attracted the economists in social science field to conduct similar research on bank data [9], national GDP [10], and sociologists on social security [11]. Subsequently, many scholars in the field of natural science have also become extremely interested in this law. It is found through statistics that computer file size [12], biological protein domain length, life distribution [13], spectrum line intensity [14] in complex atomic transition spectroscopy in physics, hadron lifetime, energy loss rate of pulsar self-rotation slowing [15] and three regular statistics, even the operation law of the objects outside the solar system in astronomy conforms to Benford’s law [16]. However, many data sets also have not passed the applicability verification, such as the resident ID number, the height of the adult. Liu Yunxia, studied the important economic data of the national development zone and found the first number, “1, 7, 8, 9”, in the tax revenue index is also inconsistent with the Benford frequency [17].

Experiments prove that the data distribution in many fields obeys Benford’s law, but there are also data sets that violate this law. Such data set can be roughly classified into two categories, one is that it does not satisfy the objective condition for applying the law; the other is that there are anomaly data, which break the normal law. In view of the former, the current theoretical basis cannot determine the reasonable data set a priori, and it still needs to be judged by rigorous statistical analysis, and this is its limitation. For the latter, scholars skillfully use this feature to develop Benford’s law as the effective tool for detecting data quality and mining anomaly data.
3.2 Anomaly data detection

Benford’s law is an effective tool for testing data quality and mining anomalous data. The principle is that if a certain data set is known to conform to the law, when a similar data set deviates, the data is likely to be abnormal.

The application of Benford’s law in the anomaly test in field of the social science is the earliest and most mature especially for artificial fraud in auditing, accounting, and financial data. Varian (1972) earlier proposed to use this law to detect whether economic data had the expiration phenomenon, but his sample data sets are fewer, which lacked persuasiveness\[18\]. Based on the system verification accounting and tax data, according to the feature that people tampered with or modified the data in line with the probability distribution of the first digit from 1 to 9 instead of the principle of logarithmic uniform distribution of the index mantissa, he proposed to use the law to detect the accounting of the enterprise ad tax data, which could not only find accounting fraud, tax evasion and tax evasion, but also could find problems beyond the data balance and made up for the shortcomings of traditional auditing methods. Subsequently, the application scope of Benford’s law was gradually extended to the financial, insurance and other industries, and it plays an important role so far\[19-21\]. In domestic research, Liu Yunxia found the quality problems of the insurance company’s “Fee and commission payment” index by means of the law, and they mainly existed in the data of “1, 4, 5, 9” in 2009-2011, because it needed to lower the data to conform to the payment limit from the regulatory authorities\[22\]. With the deepening of Benford’s law research, its application field continues to expand, and it is also good in detecting election fraud\[23\], detecting social fraud\[24\], and sports counterfeiting\[25\].

The application of Benford’s law for anomaly data detection is mainly in the social domain data. People tend to tamper with and modify data for various purposes. The natural law of digits is greatly disturbed. It is difficult to make false data conform to the law. The anomaly detection effect is very significant, but it also has the shortcomings of low precision and inaccuracy. The applications in the natural science are relatively fewer.

3.3 Cross application with other methods

The cross-application of Benford’s law with professional theories and technologies in the fields of computer, medicine, physics, astronomy and other natural sciences can optimize the actual application and solve many problems other than anomaly data detection, and the application value is extremely high. The optimization of practical application of Benford’s law first appeared in the field of computer. In the late 1980s, Barlow JL and Schatte P took the lead in combining it with computer memory and algorithm design to improve hardware performance, which reduced the memory consumption\[26\] and improved the operation speed\[27\], with the maturity of information technology, it plays a greater role in software application, for example, rapid optimization of image processing\[28\], analysis of images\[29\]; in recent years, scholars use it to detect network anomalies, and can determine whether there are intentional intrusion, unexpected failures, or general network failures\[30\] by analyzing the arrival time of SYN packet in the TCP protocol flow. In the medical field, researchers use it to track the semilethal concentration violation in animal acute toxicity tests so as to better manage toxicity data\[31\], and also propose new ideas for predicting and treating diseases\[32\].

At present, there are many cross-applications between Benford’s law and computer science, while the applications in other fields such as medicine and physics are scattered, and less progress is made. The reason for this situation is that the field is highly professional and the innovation requirements are high. Many applications are in exploration.

4. Problems in application of Benford and improvement methods

4.1 Problems in the application

4.1.1 Subject to objective data conditions
Benford’s law has been applied in many fields, but not all data conform to this law. The current theoretical basis cannot determine a priori which data is available and cannot reasonably explain its applicability\cite{33}. However, in 1997, Ni-grini summarized four broad data conditions through a series of empirical studies: the amount of data is large enough to represent all samples; the amount of the data is unlimited; the data forms naturally, with no influence of human factors or smaller influence; Data cannot be highly coupled or highly cohesive\cite{34,35}. Simply speaking, the data obtained by the measure unit system conforms to this law, and those arbitrarily obtained and restricted data usually do not conform to this law.

In practical applications, there are many data sets that cannot fully satisfy the above four conditions. If the data does not form naturally, including house number, lottery number, telephone number, date, weight, etc., such data does not apply to Benford’s law.

4.1.2 Low detection accuracy
Benford’s law is the effective tool for detecting the quality of data sets. However, as the application continues to be more diverse, the problems gradually appear and mainly are low accuracy of detection, often appearing as misjudgment of data or larger scope of the selected anomaly data pool. Through a comprehensive multi-angle analysis of the detection process and results, there are four main factors: environmental interference; data quality; single test method; system error.

(1) Environmental interference. The external environment, such as social behavior, policies and regulations or natural changes, will affect the formation of data by influencing people’s thinking and actions and resulting in the consequence of misjudgment of data. For example, for financial market data, special circumstances (such as company scandals) will affect the formation of the daily S&P500 index data, which causes the data to violate Benford’s law\cite{36}.

(2) Data quality. The data is the object of the detection and has multiple attributes, such as completeness and accuracy. The real incomplete data will cause statistics to expand. When there are more false data, they will violate the Benford’s law obviously. When there are fewer data, the overall impact will be weak and it is difficult to find anomalies.

(3) The means of detection is single. The single means of detection includes single dimension detection and single method. The former may cause misjudgment of data, and the latter may select a larger range of anomaly data pools. For example, in the United States, there is a big deviation between the statistics results based on national and state levels and those obtained according to the crime type\cite{37}, and the time-phased results are also quite different from the summary analysis results\cite{38}.

(4) System error. The significance detection of the results can improve the detection accuracy to a certain extent, but systematic errors still exist. The significance test is an important part of the analysis test results, including \(\chi^2\) goodness of fit test, KS test, distance test, Pearson correlation coefficient, etc., but for hypothesis tests of large samples and fixed level, the null hypothesis is likely to be rejected\cite{39}, type I error may also occur in other types of tests, and the abandoning trueness behavior occurs.

4.2 Improvement methods
In view of the low accuracy of test in the application of Benford’s law, the related research mainly improves in three aspects: expanding the scope of test, combining with other models, and strengthening the analysis of test results.

4.2.1 Extend detection scope
In the application of Benford’s law, it is found that the result of only analysis of the first digit is unreliable. In order to avoid the accidental error in the data verification process, it is necessary to extend the detection scope of the law from extending the data detection digits and the multi-dimensional detection data.

(1) Extend data detection digits. In general, the quality of the data set with the first digit obeying Benford’s law is good, and when the distribution probabilities of different digit combinations (such as the first- second, second-third, first- third digit, etc.) conform to the law, the test results are more
accurate. For example, when detection is carried out for some of the M2 statistical data sets issued on the official website of the People’s Bank of China, the probability distribution of the first three digits conform to Benford’s law[9], while the regression coefficient of the fabricated data is tested, the actual frequency of the second and third digit deviates from the theoretical frequency[38,40].

(2) Multi-dimensional detection data. For Benford’s law, its nature can be more fully understood by testing data from different dimensions and the shortcomings of single-dimensional analysis can be made up for. If the conclusions of several tests are consistent, the data can be considered reliable; if the results are inconsistent, there is a high probability that the data will be abnormal and further analysis is needed[38].

4.2.2 Application by combining with other models

Combining with statistical model and data mining (classification, clustering, outlier detection) model, constructing the composite test model based on Benford’s law can effectively solve the problems of data and methods existing in the application of the law, which could make great progress in the following aspects: overcoming data defects, strengthening the test methods by taking advantage of other methods and correcting test errors.

(1) Overcoming data defects. Data defects mainly include problems such as incompleteness and instability of data. When the data is incomplete, the combining of Benford’s law with unsupervised learning techniques or reinforcement learning techniques can overcome the problem of too high test probability caused by incomplete data, and can effectively detect health insurance claim application fraud[41,42]; When the data is unstable, the combination with the decision tree algorithm and distance-based method can overcome shortcomings of small scale and large change of panel data, and the accuracy rate of judging an insurance company with abnormality by using it reaches 98.66%[22].

(2) Strengthening the test methods by taking advantage of other methods. When the data has no defect, through the statistical methods such as the panel model specially, based on the distance, cluster and outlier detection algorithm, test with Benford’s law synergistically; complement each other’s advantages to make up for the deficiency caused by both of them used separately[43]. It not only can overcome the problem that the former can not determine whether the abnormal value is caused by data quality or other problems, but also can overcome the shortcomings that the latter needs to select repeatedly and cannot accurately locate the abnormal point, so that the test efficiency and accuracy are greatly improved.

(3) Correcting the test error. There are inevitable or accidental errors in the data verification process. It is an important means to improve the accuracy of the test results by adopting appropriate methods for avoiding, eliminating or correcting the errors. For example, for the problem that the null hypothesis may be rejected, Bayes factor is proposed for correcting p-value comprehensively so as to get more accurate and reliable results[39].

4.2.3 Strengthening the explanation of test results

The quality of social science data is often lower than natural science data. We can use Benford’s law and other methods to improve the efficiency of quality detection and mine anomaly data. However, we need to theoretically strengthen the analysis of test results and reduce negative impact on data quality by policies, regions, humanities, markets and various factors so as to obtain more objective and reasonable conclusions.

5. Conclusions and prospects

Benford’s law describes the distribution law of the first digit of most data in nature. It is an effective tool for testing data quality and innovative application in the fields of nature, social science and others. It helps to enhance the strategic role of data in the era of big data and support the country and various industries to make scientific decisions.

The focus of various industries is not only to judge the quality of data, but also to mine anomaly data and extract effective information. Scholars should strengthen their cooperation and efforts,
conduct in-depth research on their essence, improve the application mechanism, and fully explore their hidden potential. Firstly, the research on its basic theory should be deepened, for example, clarification of the applicable conditions of Benford’s law. Secondly, the joint application of Benford’s law with other data processing techniques should be strengthened, according to the common data quality problems in specific fields, adopting data mining techniques, traditional data detection model and other targeted optimization test methods to correct test errors and improve test accuracy. Finally, strengthen the explanation of Benford’s law results, based on the analysis of the law itself, strengthen the combination of theory with practice so as to solve more practical problems and make better use of its value.

Acknowledgements
This work was supported by the Fundamental Research Funds for Central Non-profit Scientific Institution (Grant Number JBYW-AII-2017-06, JBYW-AII-2017-33). *Corresponding author: WU Jianzhai (wujianzhai@caas.cn)

References
[1] G.J. Li, X.Q. Cheng. Big Data Research: A Major Strategic Field of Future Science and Technology and Economic and Social Development—Research Status and Scientific Thinking of Big Data. BCAS, 27(6):647-657 (2012).
[2] X.J. Tao, X.F. Hu, Y. Liu. A Review of Big Data Research. J. Syst. Simul., 25(S1): 142-146(2013).
[3] D.G. Feng, M. Zhang, H. Li. Big Data Security and Privacy Protection. Chin. J. Comput., 37(01): 246-258(2014).
[4] S. Newcomb. Note on the frequency of use the different digits in natural numbers. Am. J. Math., 4(1):39-40(1881).
[5] F. Benford. The law of anomalous numbers. Proceeding of the American Philosophical Society, 78(4):551-572(1938).
[6] T P. Hill. The significant-digit phenomenon. Am.Math.Mon, 102: 322-327(1995).
[7] T P. Hill. Base-invariance implies Benford’s law. Proceeding of the American Philosophical Society, 123: 887-895(1995).
[8] W.M. Zhu, H. Wang, W. Chen. Research on Fraud Detection Method Based on Benford’s Law. J. Appl. Stat. and Manage., (1): 41-46. (2007)
[9] D.L. Xu, Y.Jin. Research on the Accuracy of M_2 Statistical Data based on Benford’s law. SIF, (8): 20-24(2010).
[10] A .Carsten, HOLZ. The quality of China’s GDP statistics. China Econ. Rev., 30:309-338(2014).
[11] B. Rauch. Deficit versus social statistics:empirical eviense for the effectiveness of Benford’s law. Appl. Eco. Lett, 21(3):147-151(2014).
[12] J. Torres, S. Fernández, A. Gamero, et al. How do numbers begin? (The first digit law).Eur. J. Phys, 28(3):17-25(2007).
[13] L M. Leemis, B W. Schmeiser, D L. Evans. Survival distributions satisfying Benford’s law. Am. Stat., 54: 236-241(2000).
[14] J-C.Pain. Benford’s law and complex atomic spectra [J]. PHY REV E, 77: 012102(2008).
[15] L. Shao, B Q. Ma. Empirical mantissa distributions of pulsars. ASTROPAR PHYS, 33: 255-262(2010).
[16] T. Alexopoulos, S. Leontsinis. Benford’s Law in Astronomy. Astron. Astrophys., 35(4):639-648(2014).
[17] Y.X. Liu, X.M. Wu, W.Y. Zeng. Research on the Comprehensive Use of Benford’s law and Panel Model for Detecting the Quality of Statistical Data. Stat. Res., (11):74-78(2012).
[18] Varian, Hal. Benford’s Law. ASA, 26(3):65(1972).
[19] M J. Nigrini. A taxpayer compliance application of Benford’s law.JATA, 18: 72-91(1996).
[20] M.J. Nigrini, L.J. Mittermaier. The use of Benford’s law as an aid in analytical procedures. Auditing, 16: 52-67(1997).
[21] A.M. Rose, J.M. Rose. Turn excel into a financial sleuth: An easy-to-use digital analysis tool can red- flag irregularities. J. Account, 196: 58-60(2003).
[22] Y.X. Liu, W.Y. Zeng. Further Research on the Comprehensive Application of Benford’s law and Other Methods for Evaluating the Quality of Statistical Data. Stat. Res., 30(8):3-9(2013).
[23] F.Y. Lin, L.M. Guan, W.C. Fang. Heaping in Reported Earnings: Evidence from Monthly Financial Reports of Taiwanese Firms. EMERG MARK FINANC TR, 47(2), (2011).
[24] J. Golbeck. Benford’s Law Applies to Online Social Networks. Plos One, 10(8):1-9(2015).
[25] JaeHyeon. The Possibility of Detecting Match-fixing : Benford’s Law in Sports Data. The Korean Journal of Measurement and Evaluation in Physical Education and Sports Science, 19(1):69-89(2017).
[26] P. Schatte. On mantissa distributions in computing and Benford’s law. J. Inf. Process. and Cyber., 24(9): 443-455(1988).
[27] J.L. Barlow, Bareiss E.H. On roundoff error distributions in floating point and logarithmic arithmetic. Computing, 34: 325-347(1985).
[28] J.M. Jolion. Images and Benford’s law. Journal of Mathematical Imaging and Vision, 14: 73-81(2001).
[29] D.Fu, Y.Q.Sh, W.Su. A generalized Benford’s law for JPEG coefficients and its applications in image forensics. Proceedings of SPIE, 65051L-65051L-11(2007).
[30] L. Arshadi, AH. Jahangir. Benford’s Law behavior of Internet Traffic. J.Netw.Comput.Apply, 40(1):194-205(2014).
[31] Pepijn de Vries, Albertinka J.Murk. Compliance of LC50 and NOEC data with Benford’s Law. An indication of reliability. ECOTOX ENVIRON SAFE, (98): 171-178(2013).
[32] Y.K. Tang. Intrinsically disordered proteins(IDPs) Obey the Benford’s law. J. of Baoji Univ. of Arts and Sci., 34(1):1-5(2015).
[33] R. Raimi, The first digit problem, Am. Math. Mon, 83:521-538(1976).
[34] M.J. Nigrini, L.J. Mittermaier. The use of Benford’s law as an aid in analytical procedures. Auditing, 16: 52-67(1997).
[35] M.J. Nigrini. Digital Analysis Using Benford’s Law. Global Audit Publications, Vancouver, B.C., Canada,(2000).
[36] Corazza, M. Checking Financial Markets via Benford’s Law: the S&P 500 Case. Mathematical and Statistical Methods for Actuarial Sciences and Finance, 93-102(2010).
[37] MJ. Hickman, SK. Rice. Digital Analysis of Crime Statistics: Does Crime Conform to Benford’s Law? J. Antu. Criminal, (26):333-349(2010).
[38] A. Diekmann, B. Jann. Benford’s Law and Fraud Detection: Facts and Legends. Ger. Econ. Rev., (26):333-349(2010).
[39] L. Pericchi, D. Torres. Quick Anomaly Detection By the Newcomb Benford Law With Applications to Electoral Processes Data from the USA, Puerto Rico and Venezuela. Stat. Sci., 26(11):502-516(2011).
[40] A. Diekmann. Not the First Digit! Using Benford’s Law to Detect Fraudulent Scientific Data. J.Appl. Stats, 3(34):321-329(2007).
[41] F. Lu and J. E. Boritz. Detecting Fraud in Health Insurance Data: Learning to Model Incomplete Benford’s Law Distributions. In 16th European Conference on Machine Learning (pages 633–640, Porto, Portugal, 2005. Springer).
[42] F. Lu, J.E. Boritz, et al. Adaptive Fraud Detection using Benford’s Law. Advances in Artificial Intelligence, Proceedings, (4013):347-358(2006).
[43] B. Little, et al. Benford’s Law, Data Mining, and Financial Fraud: a Case Study in New York State Medicaid data. Data Mining IX: Data Mining, Protection, Detection and Other Security Technologies, (40):195-204(2008).