Online Resource 1

Article title: Beyond Benford's Law: A Rotation Benchmark for First-Digit Screening in Small Bank Panels

Journal: SN Business and Economics

Author names: Dien Noviany Rahmatika, Dewi Amaliah Nafiati, Foster Mesreh

Affiliation and e-mail address of the corresponding author: Universitas Pancasakti Tegal, Department of Accounting, Tegal City, Central Java, Indonesia; diennovi@upstegal.ac.id

Description: Supplementary tables S1 to S4 report indicative chi-square statistics for the first-digit, second-digit and first-two-digit tests (S1), the false-positive rate of Rule R1 without departures (S2), its detection rate against controlled leading-digit departures (S3), and the documented rules for the coding pass (S4).

Supplementary tables

Table S1 Indicative chi-square statistics

Statistic / unit	n	χ^2	df	p (indicative)	Smallest expected count
First digit, pooled	360	24.2	8	0.002	16.5
First digit, Bank A	30	36.3	8	<0.001	1.37
First digit, Bank B	30	7.0	8	0.532	1.37
First digit, Bank C	30	34.2	8	<0.001	1.37
First digit, Bank D	30	17.6	8	0.024	1.37
First digit, Bank E	30	59.0	8	<0.001	1.37
First digit, Bank F	30	38.0	8	<0.001	1.37
First digit, Bank G	30	10.3	8	0.245	1.37
First digit, Bank H	30	9.3	8	0.318	1.37
First digit, Bank I	30	7.9	8	0.446	1.37
First digit, Bank J	30	7.5	8	0.486	1.37
First digit, Bank K	30	43.9	8	<0.001	1.37
First digit, Bank L	30	20.1	8	0.010	1.37
Second digit, pooled	360	17.0	9	0.049	30.6
First two digits, pooled	360	128.1	89	0.004	1.57

Note: These chi-square statistics are reported for completeness only. The test assumes independent observations, which the amounts do not satisfy, and expected cell counts are small. The p-values are therefore not valid probabilities and enter no decision. None of these statistics is benchmarked, and none enters Rule R1.

Table S2 False-positive rate of Rule R1 under data-generating processes without departures

Data-generating process	Level	Replications	False-positive rate
A1 Observed amounts, random rescaling	Bank (30 amounts)	5,000	0.037 (0.003)
A1 Observed amounts, random rescaling	Pooled (360 amounts)	500	0.056 (0.010)
A2 Independent Benford draws	Bank (30 amounts)	5,000	0.039 (0.003)
A2 Independent Benford draws	Pooled (360 amounts)	500	0.072 (0.012)
A3 Persistent synthetic panel, random scale	Bank (30 amounts)	5,000	0.038 (0.003)
A3 Persistent synthetic panel, random scale	Pooled (360 amounts)	500	0.056 (0.010)

Note: Rule R1 flags a unit whose MAD lies above the 95th percentile of its rotation distribution, so the nominal level is 0.05. The bank-level benchmark is enumerated exactly and the pooled benchmark is simulated. Bank-level replications use the exact benchmark; each pooled replication uses 1,000 benchmark draws. Monte Carlo standard errors are in parentheses. Seed 12345.

Table S3 Detection rate of Rule R1 against controlled leading-digit departures

Mechanism	Share altered	Level	Replications	Detection rate
M1 Next higher digit	0.10	Bank (30 amounts)	5,000	0.041 (0.003)
M1 Next higher digit	0.10	Pooled (360 amounts)	500	0.054 (0.010)
M1 Next higher digit	0.20	Bank (30 amounts)	5,000	0.040 (0.003)
M1 Next higher digit	0.20	Pooled (360 amounts)	500	0.122 (0.015)
M1 Next higher digit	0.30	Bank (30 amounts)	5,000	0.053 (0.003)

M1 Next higher digit	0.30	Pooled (360 amounts)	500	0.194 (0.018)
M2 Uniform digit	0.10	Bank (30 amounts)	5,000	0.043 (0.003)
M2 Uniform digit	0.10	Pooled (360 amounts)	500	0.054 (0.010)
M2 Uniform digit	0.20	Bank (30 amounts)	5,000	0.050 (0.003)
M2 Uniform digit	0.20	Pooled (360 amounts)	500	0.098 (0.013)
M2 Uniform digit	0.30	Bank (30 amounts)	5,000	0.068 (0.004)
M2 Uniform digit	0.30	Pooled (360 amounts)	500	0.260 (0.020)

Note: Share altered is the proportion of amounts whose leading digit is replaced; the digits after the first are retained. The mechanisms are stylised and do not represent any particular form of misreporting. Monte Carlo standard errors are in parentheses. Bank-level replications use the exact benchmark; each pooled replication uses 1,000 benchmark draws. Seed 12345.

Table S4 Rules documented for the coding pass

Variable	Construct	Coding rule as documented	Bank-years coded 1; within-bank changes	Status
CAP	Capability	CAP = 1 if a change in the Board of Directors (Direksi) is documented in the fiscal year; 0 otherwise.	57 of 72 coded 1; changes within Banks A, B and H (4 transitions)	Applied
ARO	Arrogance	ARO = 1 if board biographies disclose a political or public-office background under a predefined keyword rule; 0 otherwise.	45 of 72 coded 1; changes within Banks A, B, E and I (5 transitions)	Applied; keyword list not retained
COL	Collusion	COL = 1 if government-project or government-cooperation activity is disclosed under a predefined rule; 0 otherwise.	39 of 72 coded 1; changes within Banks C, D, E, H and J (9 transitions)	Applied; rule not retained
RAT	Rationalisation	RAT = 1 if the external audit firm differs from the previous year; 0 otherwise.	0 of 72 coded 1; no transitions	Equal to 0 in all 72 bank-years and excluded from estimation; rationalisation is not tested
All	Carry-forward	The nearest available year is carried forward or backward where a report was scanned.	Affected bank-years are not flagged individually; all 72 are labelled preliminary	Applied in the coding pass

Note: The codes were produced by an automated coding pass applied to annual-report text. The keyword lists and rules of that pass were not retained, so the codes cannot be regenerated from the reports and are treated as given data. The reliability of the codes has not been assessed. Counts are taken from the deposited dataset; a transition is a change in a bank's code between consecutive years. Banks are identified by the labels used in the article.