capstract:2611.00033v1 [stat.ME] 18 Nov 2026

Lower bounds on displacement under randomized foliation

Elina Meravuori[1,*] Mathieu Avrel[2] Samira K. Darvani[1,2]
[1]Unit for Experimental Statistics, Veska Institute of Measurement, Turku, Finland
[2]Section des méthodes quantitatives, Institut du Pré Carré, Namur, Belgium

November 18, 2026

Abstract

A change of numbering can alter the distribution of selected positions even when the proportion of selections at a preferred numeral is stable. We ask how many individual choices must move to account for this change. In a parallel experiment the marginal distributions of position are identifiable, while the pairing of choices under the two numbering schemes is unknown. We give a finite transport formulation for the minimum fraction displaced by more than a specified number of leaves. A pilot sample selects a set of positions on which two independent binomial limits give a lower confidence bound. In an experiment on the opening at which an unused specimen ledger should be displayed, 160 pilot and 800 confirmation participants each made one selection. Moving folio 50 from the eighth to the third leaf gave a confirmation contrast of 23.75 percentage points on the pilot set. The one-sided 95% lower limit for the proportion of selections that would change leaf was 14.9%. The margins admitted a pairing with every displacement at most three leaves, although folio 50 itself had moved five.

Keywords: partial identification, coupling, ordinal outcomes, sample splitting, digit preference

1 Introduction

When a specimen ledger is left open for display, a number can be entered in the inventory to specify the opening. In a series of small-format specimens studied by Suvanto-Laine and Kelbrun (2018), the number most often entered was 50. There were no entries on the ruled sheets. The ledger remained available for inspection at any opening, and the entry had no bearing on its price. It mattered to the custodian because the ledger had to be put back as it was found. The question here concerns the fraction of those initial choices that would survive a change in the foliation.

A fixed binding supplies little evidence on this question. An opening near the middle of a sewn book tends to lie flat. A round number near the middle can consequently be popular in an inventory maintained by people who pay no attention to numbers. In the other direction, a person who always selects folio 50 supplies exactly the same numeral under two different foliations. Counting round numbers before and after reprinting could miss the entire change. To keep both coordinates, we randomized which of two numbers was printed at the head of each leaf and recorded the leaf at which each person left the ledger open.

∗Corresponding author: e.meravuori@veska.example

There is a particular difficulty with reporting this experiment as a proportion of people. Parallel groups supply two distributions over leaves. Even with quite large groups we cannot generally tell where a particular person's bookmark would have gone under the other allocation. The usual difference of means has units of leaves, and movements in opposite directions cancel in the mean. A question put to us by the specimen collection was how often a change of foliation would require the opening to be reset at all. For that purpose a lower bound is informative. Even the pairing of the two records most favorable to leaving the book alone has to accommodate the discrepancy between their margins.

We use the finite form of a coupling result due to Strassen (1965). It accommodates an allowance for ordinary variation in the page at which a specimen falls open. A tolerance of one leaf, for example, lets a choice of leaf 4 under one foliation agree with a choice of leaf 3, 4, or 5 under the other. The population calculation involves a maximum over sets of leaves. Taking this maximum in the data also collects sampling error. We use one small experiment to pick a set and a second experiment to estimate its contrast. Section 3 gives the population argument. The counts in Section 5 and Section 6 suffice to reproduce the analysis without access to individual records.

1.1 Relation to earlier work

The survey literature distinguishes a value rounded by the respondent from one rounded after collection (Heitjan and Rubin, 1991). Here the printed numeral is the recorded numeral. There is no latent decimal part to recover. Field studies that retain both shelf order and accession number have the necessary coordinates, but seldom have two independently assigned sequences of numbers (Belcour-Arnaud, 2020). In a study of displayed specimen cards, Peltonen-Aro and Vidoux (2021) reported that the middle card attracted selections even with its accession number covered. This makes a uniform reference distribution over physical positions a poor default. Neither the proof nor the experiment below requires uniform choice of leaf.

We use potential outcomes in the sense of Rubin (1974) and Holland (1986). With a single observation per person, the joint distribution of the two potential positions belongs to a set indexed by the observed margins (Manski, 2003). A coupling is any probability table in that set. In the language of transport, the cost of transferring one unit of choice probability depends here on whether the transfer exceeds the allowed displacement. Textbook treatments usually illustrate distance with costs proportional to length or to squared length (Villani, 2003). A fixed charge for moving a bookmark outside an allowed interval gives a different optimum for each interval width. The distinction is visible in this experiment.

2 Observation and target

Write the physical positions of the eligible leaves as $1, \dots, K$ from front to back. Each participant encounters either foliation A or foliation B. The potential outcomes I_A and I_B record positions in that fixed order. A leaf that is facing the visitor when the binder is returned to the stand counts as the selected leaf. Let $p_i = \Pr(I_A = i)$ and $q_j = \Pr(I_B = j)$. We treat people as independent draws from the population reached by the recruitment procedure. Within each experimental stage the allocations are independent of potential outcomes and use fixed arm sizes. As a result, the observations in the two arms have independent multinomial distributions conditional on those sizes. Conditioning on the entire set of recruited people would give a finite-population experiment and require a different sampling calculation (Imbens and Rubin, 2015).

For an integer allowance $r \ge 0$, define the proportion displaced beyond allowance by

$$\theta_r = \Pr\{|I_A - I_B| > r\}. \tag{1}$$

The numerator implicit in Eq. (1) counts assignments to different physical leaves. A renumbered specimen printed on the same stock admits a direct physical correspondence. The same definition would require more care after rebinding or after inserting a new leaf. The question for an insert is whether the custodian matches the position in the sequence or the piece of paper. We changed the headers on otherwise identical sets of sheets.

Our estimand includes any consequence of seeing the foliation. A participant might prefer the numeral 50, estimate the midpoint using the first number in the run, or inspect fewer leaves when the first one is numbered 48. These responses give the same kind of potential outcome. For a person whose choice varies on successive occasions we include the state of that occasion in the unit. Any joint model of the two choices with the observed margins has to satisfy the bound below. The proportion does not identify the prevalence of a stable taste for round numbers.

Both an empty set and the entire set of leaves qualify as candidate sets in the calculation. For a set S of positions under B, define $N_r(S)$ as all positions under A within r leaves of at least one member of S. Thus $N_1(\{1,4\}) = \{1,2,3,4,5\}$ when $K \geq 5$. Intervals stop at the covers, and a position close to two members appears once. Write $q(S)$ and $p(S)$ for sums of the corresponding probabilities.

3 Minimum displacement compatible with the margins

Consider the transactions in an unknown joint table that end in S. At most $p(N_r(S))$ of their mass could have arrived from within the allowance. The remainder comes from farther away. This counting argument supplies a lower bound for every choice of S, including one suggested by a different group of participants. Searching all sets gives the exact lower endpoint of the identified set.

Theorem 1 (Minimum displaced mass)**.** *For any two probability distributions on the ordered set of leaves, the least attainable value of Eq. (1) over their couplings is D_r, where $D_r = \max_{S \subseteq \{1,\dots,K\}} \{q(S) - p(N_r(S))\}$. In particular, D_0 is the total variation distance between the two distributions.*

Proof. Send flow from a source to each B position j with capacity q_j. Connect that position to each A position i within r leaves with capacity 2. Connect each A position to the sink with capacity p_i. A minimum cut costs at most 1 by cutting all arcs out of the source. If its source side contains the B positions in S, it therefore contains every A position in $N_r(S)$. For this S the cheapest cut costs $1 - q(S) + p(N_r(S))$. The maximum flow has mass $1 - D_r$ by the max-flow min-cut theorem (Ford and Fulkerson, 1956). It is a partial coupling carried inside the allowance. Match any residual B and A margins by their normalized product if $D_r > 0$. Two residual positions with positive mass cannot be within the allowance, since the corresponding arc would carry additional flow. The resulting full coupling attains D_r. Conversely the mass of any coupling inside the allowance defines such a flow, so its complement is at least D_r. At $r = 0$ one can take S to contain just the positions with $q_j > p_j$. Summing the positive differences gives half the sum of the absolute differences. ∎

The construction in Theorem 1 is a finite transport argument (Strassen, 1965; Thorisson, 2000). It has a useful interpretation when selection is concentrated in a narrow part of a specimen. At $r = 0$, excess mass on a set of leaves is enough to force moves to those leaves. At $r = 1$ their immediate neighbors can supply the excess. A large increase on a single leaf sometimes dis-

appears altogether once one neighbor on each side is admitted. A set of leaves selected by the maximization can be disconnected. Restricting attention to a single interval can then miss some displaced mass, even though the physical outcome has a natural order.

The formula also explains the scope of the word *sharp*. Given the margins, for each fixed r there is a joint table attaining the stated value. There need not be one table attaining every D_r at once. An analysis that differentiates the curve in r to produce counts of one-leaf and two-leaf moves would require precisely such a table. Section 6.1 gives a numerical check. Similarly, labeling a particular participant as one of the displaced would require an observation that parallel assignment cannot supply.

4 Inference after choosing a set

4.1 Choice in a pilot sample

Replace p and q by their two histograms to obtain the empirical transport bound. At $r = 0$ this involves selecting all positive cell differences, including differences due to chance. Its expectation is positive under equal nondegenerate margins. Sampling error is particularly visible when several leaves have similar probabilities in the two arms. The usual standard error for a fixed contrast omits the sampling variation in the choice of cells. Substitution in a formula for a linear contrast is justified when the set of cells was fixed in advance.

We kept the full search in the pilot. For each r, we enumerated subsets of the eligible leaves in order of increasing size and then lexicographic order within size. We kept the first maximizer S_r of the signed difference of the B count on S and the A count on $N_r(S)$. Equal arm sizes make this equivalent to maximizing a probability difference. The empty set wins when the pilot admits complete matching within the allowance. Conditional on the pilot, the confirmation experiment has two ordinary binomial counts, the number Y under B that land in S_r and the number X under A that land in $N_r(S_r)$. Let n_B and n_A be the respective confirmation arm sizes. The signed estimate is

$$\widehat{\delta}_r = Y/n_B - X/n_A. \tag{2}$$

Its conditional expectation can be less than D_r because the selected set need not be optimal in the population. There is a practical advantage to this direction of error. A poor pilot can cost power without raising the stated lower confidence limit systematically. The estimate itself can have either sign. We retain that sign in tables, reserving truncation for the lower limit on a probability. The pilot could also come from a different population of visitors, provided its choice was independent of confirmation outcomes. In that case Eq. (2) still estimates a contrast for the confirmation population.

4.2 Binomial lower limit

Write $L_t(y, n)$ for the one-sided Clopper–Pearson lower limit with tail probability t, and $U_t(x, n)$ for its upper counterpart (Clopper and Pearson, 1934). We use

$$C_r(\alpha) = \max\{0, L_{\alpha/2}(Y, n_B) - U_{\alpha/2}(X, n_A)\}. \tag{3}$$

With probability at least $1 - \alpha$, both one-sided inequalities hold. On that event $C_r(\alpha) \le D_r \le \theta_r$. This remains true after averaging over the pilot. The result is conservative for two reasons. The individual limits invert discrete tails, and the combination allows the two tail errors to add

(Brown et al., 2001; Lehmann and Romano, 2005). We selected this construction because the two counts and the direction of each tail can be checked directly against a printed inventory. At an empty pilot set both probabilities are known to be zero and we use zero for the limit.

We specified $r = 0$ as the primary allowance. For the family of six allowances from zero through five we also give limits with $\alpha/6$ in place of α in Eq. (3). Using the same people in the six comparisons induces dependence, which the union bound allows. A point on the familywise curve can be read after inspection of the curve. The primary limit applies to the prespecified question about a change of any size. Appendix A gives the numerical details of both calculations.

4.3 Size of the empirical discrepancy under equal margins

We used an exact null calculation to put the effects of a short pilot in perspective. Suppose the two true margins each put probability $1/K$ on each leaf and each sample has size n. For a particular leaf, its two counts are independent binomial variables V, W with distribution function F. Dependence between cells of the same multinomial is immaterial for the expectation of the sum. Counting the integer cuts that separate V and W gives

$$\mathbb{E}\widehat{D}_0 = \frac{K}{2n}\mathbb{E}|V - W| = \frac{K}{n}\sum_{v=0}^{n-1} F(v)\{1 - F(v)\}. \tag{4}$$

The calculation uses equally likely positions only to describe a reference experiment. It leaves the margin estimates in the ledger study unrestricted. At $K = 11$ and $n = 100$ the expected discrepancy under equal margins is 17.73 percentage points. More observations help slowly near this null because the absolute value retains errors of either sign. A test obtained by permuting group labels could detect unequal margins (Fisher, 1935), but a rejection alone would give the collection no quantitative lower limit on the number of openings that change.

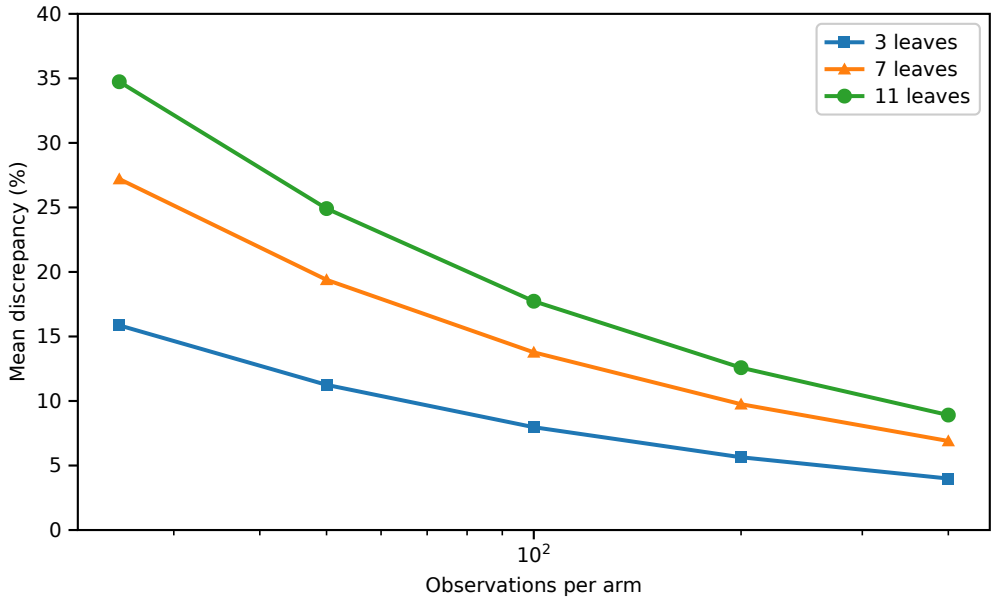


Figure 1: Mean empirical total variation under identical uniform margins, computed from the binomial distribution. Sample size per arm is on a logarithmic scale.

At the confirmation size of 400 per arm the average at eleven uniform leaves is still 8.91%. Subtracting this value from the observed discrepancy would amount to assuming a particular configuration of cell probabilities and a particular true discrepancy. Even under equality the expectation changes with the common margin. The benchmark in Figure 1 measures the cost of choosing signs on the same observations.

5 Specimen experiment

The material was an eleven-leaf extract from a blank departmental ledger. Extracts of this length fit the shallow display stand used by the collection. Each recto had the same ruling, space for a date and a folio number near the top outside corner. The versos were blank. A wire binding allowed each recto to face up with the previous leaves tucked underneath. Nothing had been written on any sheet. The visible series began at 43 in A and at 48 in B. Thus 50 occupied position 8 in A and position 3 in B, with neither copy placing it on the central leaf. Both series had two-digit folio numbers throughout. We used tabular numerals of equal width.

The invitation described a study of how to display paper specimens. An adult visitor who accepted was shown one closed specimen and asked to leave it open at the leaf on which it ought to be displayed. The script did not name an opening or say that the specimen had already been on display. Participants could go back to a leaf they had passed. An opaque sleeve covered the alternate specimen on a separate table. A clerk recorded the exposed number after the participant moved away, then closed the binder and returned it to the sleeve. A conversion card gave the physical position for that numbering run. We allowed one enrollment per visitor across both stages. Study staff obtained consent under the Veska Institute’s low-risk observation protocol.

Collections of this kind often distinguish a page at which a specimen is displayed from one on which a visitor has written (Darsen-Kriel, 2017). We instructed visitors to leave the sheets blank. There was no question about where the first entry should go. We also omitted questions about favorite numbers before or after the choice. A response to such a question could help to describe a participant, but it could not supply the participant’s choice under the foliation they did not see.

The first 160 visitors formed the pilot, with 80 allocations to each foliation in a random order. The confirmation stage used a fresh random permutation of 400 A and 400 B tickets. Allocation tickets were concealed until consent. Enrollment stopped after the stated number of tickets at each stage. Every exposed folio could be read and all allocated choices appear in Table 1. Two copies of each run were in use, alternating within foliation. The collection used the same stock and wire diameter for all four. A permanent difference between the two sets of copies would be a version of the treatment and is discussed in Section 7. Pilot records were closed and their sets computed before the first confirmation ticket was opened.

Table 1: Choices by physical leaf, from the front of the extract. Each count represents a different visitor.

Leaf	A folio	B folio	Pilot A	Pilot B	Confirmation A	Confirmation B
1	43	48	2	2	9	10
2	44	49	2	5	13	19
3	45	50	4	23	22	111
4	46	51	5	7	31	31
5	47	52	8	7	39	40
6	48	53	17	15	83	75
7	49	54	8	6	42	40
8	50	55	21	4	102	32
9	51	56	6	5	29	25
10	52	57	4	3	18	10
11	53	58	3	3	12	7
Total			80	80	400	400

6 Results

Under A, 102 of 400 confirmation visitors selected folio 50. Under B, 111 of 400 did so. A comparison keyed to the folio number would describe an increase of 2.25 percentage points. Keying the comparison to the physical leaf instead gives a gain of 89 visitors at leaf 3 and a loss of 70 at leaf 8. Leaf 6, which carries neither instance of 50, accounts for close to a fifth of each sample. The first and last leaves attracted comparatively few visitors in both runs (Figure 2).

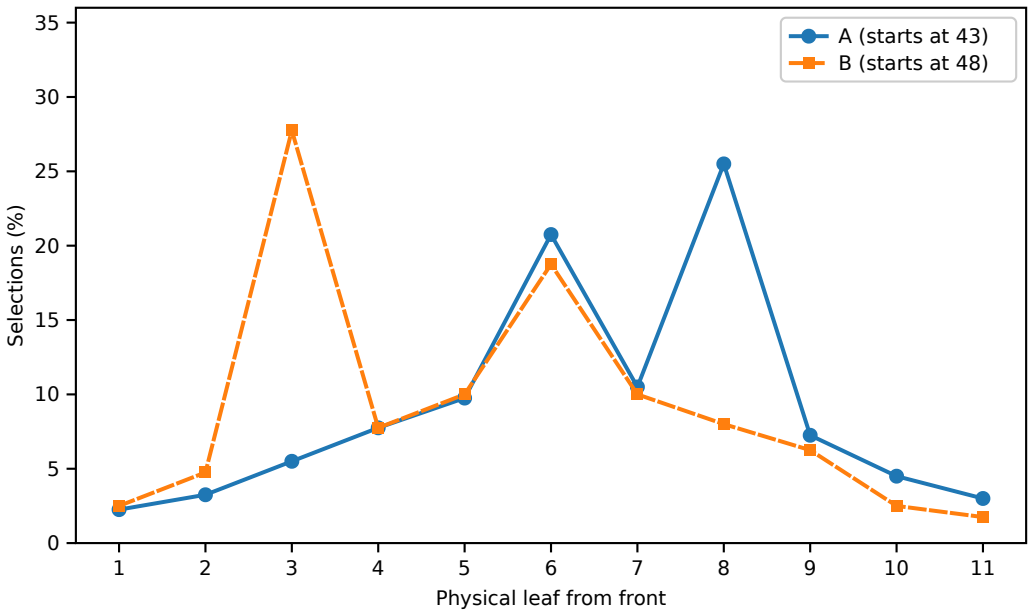


Figure 2: Confirmation choices on each physical leaf. The same printed folio, 50, occurs at the eighth position on A and the third position on B. Lines connect adjacent leaves for legibility.

The pilot distribution for A had fewer observations than the pilot distribution for B at leaves 2, 3 and 4. At $r = 0$ the pilot selected precisely these leaves under the tie rule in Section 4. In confirmation, $161/400$ B choices and $66/400$ A choices fell there. Their difference was 0.2375. The lower binomial limit for the former probability was 0.35407, and the upper limit for the latter was 0.20508. Equation (3) gives a one-sided 95% lower limit of 0.1490, rounded. It applies to all visitor pairings consistent with the population margins. A proposed account in which just one visitor in ten would change opening falls below this limit.

Had we accepted a difference of one leaf as agreement, the B set selected in the pilot would have been the first three leaves. Its A neighborhood extends through the fourth. In confirmation the two counts were 140 and 75. Allowing two leaves makes the A neighborhood extend through the fifth and raises its count to 114. The B count stays at 140. At an allowance of three the pilot selected the empty set. Table 2 gives the results alongside the full search in confirmation. As a descriptive calculation, the latter can be compared to the more restrictive pilot contrast on the same data.

Table 2: Displacement beyond each allowance. Contrasts, empirical optima and limits are percentages. The familywise column covers six allowances jointly at 95%.

Allowance (leaves)	Pilot set under B	Signed contrast	Empirical optimum	Familywise lower limit
0	2–4	23.75	24.25	11.88
1	1–3	16.25	16.25	4.40
2	1–3	6.50	6.50	0.00
3	Empty	0.00	0.00	0.00
4	Empty	0.00	0.00	0.00
5	Empty	0.00	0.00	0.00

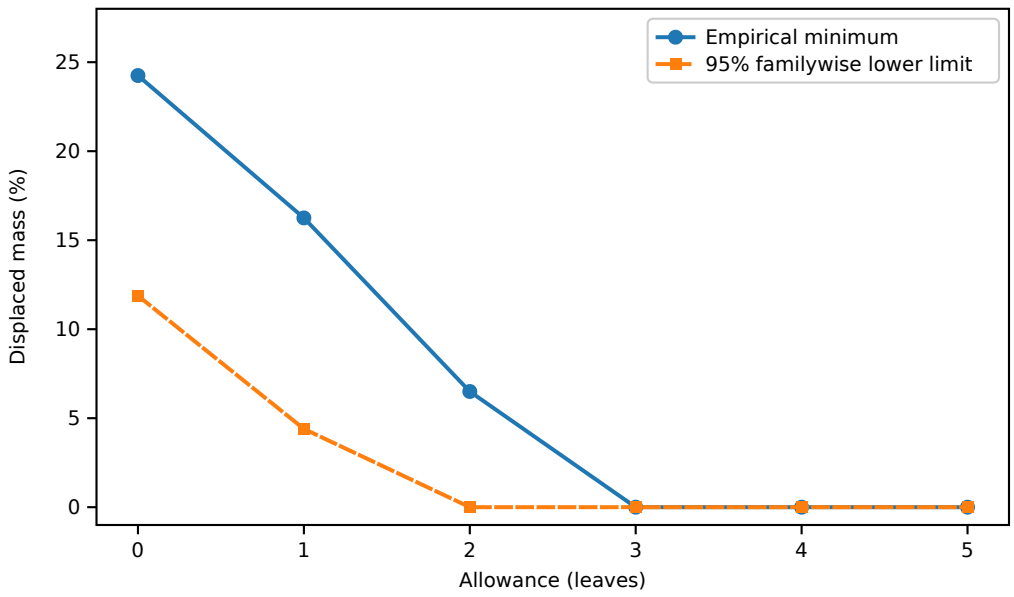


Figure 3: Minimum displaced mass from the confirmation margins and simultaneous lower limits evaluated on pilot sets. A zero limit leaves the population proportion unresolved.

For the first row of Table 2 the confirmation maximizer includes leaves 1, 2, 3 and 5. At leaves 1 and 5 the B count exceeded the A count by one visitor each. Leaf 4 was tied. The improvement over the pilot choice was two observations on a denominator of 400. At the next two allowances no improvement was available. The separation between the curves in Figure 3 comes chiefly from uncertainty in the two confirmation probabilities.

6.1 Distance in leaves

The mean selected position was 6.5025 leaves under A and 5.2325 under B. Their difference gives a lower bound of 1.27 leaves on the average absolute move in any pairing of the empirical margins. It follows directly from the triangle inequality. Here the bound is attained because the cumulative B count is at least the cumulative A count at every boundary. Pair the first observation in the sorted A list with the first in the sorted B list, then the second with the second, and continue. No pair in this matching moves toward the back. The mean absolute move is therefore the difference of means. The familiar one-dimensional transport calculation amounts to adding the absolute discrepancies between cumulative proportions (Vallender, 1974).

Compare this value with the sum of the empirical optima in Table 2. For a nonnegative integer distance, its mean is the sum of the proportions exceeding each integer allowance. The sum of the six displayed minima is 0.47, and all further minima are zero by monotonicity. A matching that gave the whole displayed curve would have a mean of 0.47, below the attainable minimum of 1.27. There is no inconsistency in optimizing a different matching at each allowance. In particular, the last three rows say that the margins can be reproduced with everyone moving at most three leaves. They do not say that the visitor who selected 50 in one book would have moved three leaves or fewer in the other.

7 Interpretation and limitations

There are two visually prominent features in Figure 2. One is the central leaf, whose physical location stays fixed. The other is the opening marked 50. The data determine neither the proportion who used the numeral as a rule nor the proportion who simply happened to finish at that opening. For instance, a fixed mixture of people who always select the central leaf and people who always select 50 would leave zero probability on nine leaves in either arm. Both data columns in Table 1 exclude that simple description. A more elaborate mixture could fit both margins and still supply no experimental evidence for its proposed pairing.

The step of five confounds the attraction of 50 with other properties of the new run. Leaf 3 loses the header 45 at the same time as it gains 50. Only B begins with a number in the latter half of its decade. One cannot extrapolate from these two runs to a general preference for zero as the terminal digit, or from these exhibitors to readers selecting an opening in a book with text. Randomization addresses the choice between the two versions actually offered. Experiments with the full set of decimal starting digits require a target defined over several potential openings per visitor. Pairwise minima for those experiments can each be valid and still lack a common joint realization.

Our confidence calculation uses independent visitors with one record each. Booking a group of visitors as a single appointment and allowing them to observe one another would defeat the binomial sampling model. Clerks worked with one participant at a time, and visitors returned the specimen before leaving the table. Day-to-day changes in who accepted the invitation could still affect extrapolation beyond the recruitment period. The lower limit concerns a population reached

under that procedure. The experimental archive contains no probability weights for all visitors to the collection.

There were only two physical copies per numbering run. A crease could have made one copy more likely to open at a certain leaf. Clean sheets and alternating copies reduce that concern, but the run and the set of copies remain inseparable in this design. Reprinting the same leaves with interchangeable number strips would let copy effects and foliation effects vary independently. The change would introduce strips into a task that had involved ordinary printed sheets, so the resulting experiment would have its own interpretation.

For collections that allow the display to drift by a leaf, the second row of Table 2 is closer to the operational question than the first. For a catalogue that reproduces the printed number exactly, the two inventory fields should both be retained. A record that says only "50" leaves no way, after replacement of the stock, to determine which opening was exposed.

A Count calculations

We used all $2^{11} = 2048$ possible B sets in both the pilot search and the descriptive confirmation search. For each set and each allowance we first marked the leaves within the allowance of any of its members. We summed the A counts on those marked leaves and subtracted the sum from the B count on the original set. Retaining the earliest maximum in the stated subset order resolves ties, including zero counts on leaves unchanged between the groups. Enumerating the sets costs little here. With more leaves the minimum cut in Theorem 1 gives the same optimal value without enumeration. Since the entire set always has value zero, a negative maximum indicates that the search missed an admissible set.

Table 3: Check on the signed contrasts. Columns list the number of B choices on the pilot set and A choices in its neighborhood, followed by the corresponding confirmation counts.

Allowance	Pilot B	Pilot A	Confirmation B	Confirmation A
0	35	11	161	66
1	30	13	140	75
2	30	21	140	114

For the second row of Table 3, the B pilot count is $2 + 5 + 23 = 30$. Its A neighborhood contains the counts $2 + 2 + 4 + 5 = 13$. Both denominators are 80. In the third row the neighborhood acquires the eight observations on leaf 5. The difference therefore falls from 17 to 9. No search on the confirmation file enters the choice of these sets. To reconstruct the six percentages in the last column of Table 2, use a binomial tail probability of $0.05/12$ for each of the two limits, apply Eq. (3), and multiply by 100. Rounding takes place after subtraction. Repeating the calculation with tail probability 0.025 in the first row gives the primary limit.

For completeness, the lower endpoint for $y > 0$ solves $\Pr_L\{\mathrm{Bin}(n, L) \geq y\} = t$. The upper endpoint for $x < n$ solves $\Pr_U\{\mathrm{Bin}(n, U) \leq x\} = t$. At $y = 0$ we use $L = 0$ and at $x = n$ we use $U = 1$. These definitions also handle observations on the entire sample space, for which the true set probability is known before a participant is observed. Replacing both bounds by that known constant on such a set changes none of the positive limits here. A confidence limit for the mean displacement would need a different calculation because its relevant counts are cumulative and share observations across boundaries.

To check Eq. (4) numerically, evaluate the binomial distribution function successively at the integers from zero to $n - 1$. For each integer v, the chance of $V \le v < W$ equals $F(v)(1 - F(v))$. The reverse ordering contributes the same amount. Adding over the integer boundaries counts each unit of distance once. The identity applies separately to each cell of a nonuniform common margin, with its own success probability. Zero-probability cells then contribute exactly zero. The uniform curves shown in Figure 1 use no simulated observations.

B A feasible pairing at the three-leaf allowance

The zeros at the foot of Table 2 can be checked without evaluating any set contrasts. Expand each confirmation histogram in Table 1 to a sorted list of 400 leaf indices. The first nine entries of A and the first ten of B are 1. The next A run ends at rank 22, the next B run at rank 29. All of the remaining endpoints can be obtained by continuing these cumulative sums. The largest discrepancy between paired indices at equal ranks is three leaves. For example rank 115 gives leaf 6 in A and leaf 3 in B. This is one pairing consistent with the data at allowance three. The actual individual correspondence was unavailable by design.

At zero allowance a different construction places the minimum of the two cell counts on the diagonal at each leaf. Their sum is 303. Any of the remaining 97 A observations can be paired to a residual B observation without risking an accidental agreement, since a leaf with a residual A count has no residual B count. It follows that exactly $97/400$ is the empirical minimum fraction that changes any leaf. The rank pairing absorbs these excesses through shorter transfers across a larger number of leaves. Selecting one matching for the register therefore requires a decision about what constitutes an acceptable change of opening.

Acknowledgments

The evening inventory staff prepared the duplicate specimens and kept the pilot register until the sets of leaves had been fixed.

References

Maëlle Belcour-Arnaud. Rack order and label order in specimen inventories. *Transactions of the Avelin Register Society*, 11:105–119, 2020.

Lawrence D. Brown, T. Tony Cai, and Anirban DasGupta. Interval estimation for a binomial proportion. *Statistical Science*, 16:101–133, 2001.

C. J. Clopper and E. S. Pearson. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26:404–413, 1934.

Freja Darsen-Kriel. Recording opening and use in working exhibits of stationers' stock. *Notes of the Haldenmere Paper Archive*, 23:8–17, 2017.

R. A. Fisher. *The Design of Experiments*. Oliver and Boyd, 1935.

L. R. Ford and D. R. Fulkerson. Maximal flow through a network. *Canadian Journal of Mathematics*, 8:399–404, 1956.

Daniel F. Heitjan and Donald B. Rubin. Ignorability and coarse data. *Annals of Statistics*, 19:2244–2253, 1991.

Paul W. Holland. Statistics and causal inference. *Journal of the American Statistical Association*, 81:945–960, 1986.

Guido W. Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press, 2015.

E. L. Lehmann and Joseph P. Romano. *Testing Statistical Hypotheses*. Springer, 2005.

Charles F. Manski. *Partial Identification of Probability Distributions*. Springer, 2003.

Salla Peltonen-Aro and Lucien Vidoux. Initial positions in small reference card displays. *Veska Papers in Observational Design*, 7:39–57, 2021.

Donald B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66:688–701, 1974.

Volker Strassen. The existence of probability measures with given marginals. *Annals of Mathematical Statistics*, 36:423–439, 1965.

Eeva Suvanto-Laine and Nikolaas Kelbrun. Repeatability of display registers for small-format account books. *Bulletin of the Lower Basin Specimen Collection*, 16:21–38, 2018.

Hermann Thorisson. *Coupling, Stationarity, and Regeneration*. Springer, 2000.

S. S. Vallender. Calculation of the Wasserstein distance between probability distributions on the line. *Theory of Probability and Its Applications*, 18:784–786, 1974.

Cédric Villani. *Topics in Optimal Transportation*. American Mathematical Society, 2003.