Proof of Cash as a Conservation Law: Reconciliation, Matching, and the Limits of Zero Variance

Daniel Ray Edgar
Finsider Labs

October 2026

Abstract

A proof of cash sets the ledger's account of cash against the bank's account of the same cash and requires the two to agree up to timing. I treat it as what it is mathematically, a conservation law checked against an external record, and I state both its power and its blind spots exactly. Starting from an item-level model of the two records, I derive the reconciliation identity and the four-column proof of cash, and I show that the residual in each flow column equals the net unmatched activity of the period, so that the procedure tests flows and not merely two endpoint balances. I then prove four limits. A single aggregate variance is blind to an error subspace of dimension $n-1$, and k tie-outs are blind to one of dimension at least $n-k$. A variance computed on observed account-periods is invariant to anything in unobserved ones, so no bound on total misstatement exists without a coverage ratio. Internal transfers cancel in net movement and inflate both gross columns, with a one-sided residual when a leg falls outside the period or the covered set. Matching bank lines to ledger lines is an assignment problem when one to one and contains subset sum when one to many, and the expected number of coincidental exact-amount pairs is nmp_c, linear in the date window. Finally I derive the bridge from cash received to revenue. A zero variance is necessary for clean books and far from sufficient; I close with four falsifiable predictions and the limits of the method.

1 Introduction

Every accounting system since Pacioli [1494] rests on a conservation law. A transaction is a list of signed postings that sums to zero, so value is neither created nor destroyed by the act of recording it. The algebraic treatments of Ijiri [1967] and Ellerman [1985] make this precise: the books are a vector of account balances, each transaction is a vector whose components sum to zero, and the trial balance is the statement that a sum of zero-sum vectors is zero-sum. That law is internal: it says the books are consistent with themselves, not with the world.

Cash is the one account for which the world keeps its own copy: the bank maintains an independent record of the same balance and the same flows. Auditing standards rank evidence obtained from independent external sources above evidence generated inside the

entity [International Auditing and Assurance Standards Board, 2009b, Public Company Accounting Oversight Board, 2010], and the bank statement is the leading example. A proof of cash checks the internal conservation law for the cash account, balance by balance and flow by flow, against that external record.

In M&A financial diligence every adjustment to reported earnings is an argument, and the proof of cash is where the argument touches something a counterparty cannot easily restate. A "zero variance" is therefore often presented as if it settled whether the books are right. This paper is about exactly how much it settles.

Two opposite errors are available. The first is to treat a tied proof of cash as a certificate of completeness: offsetting errors, uncovered accounts and uneliminated transfers all survive a zero variance (Sections 3 and 4). The second is to treat the procedure as a formality: the four-column form detects errors that no pair of balance reconciliations can see (Section 2). The honest position sits between, and it can be stated as theorems.

Section 2 proves the reconciliation identity and the four-column theorem; Section 3 the kernel and coverage theorems; Section 4 the transfer formulas; Section 5 treats matching as optimisation; Section 6 derives the bridge to revenue. A worked example, implications, predictions and limitations follow. Nothing here is an empirical result. Every numerical example is illustrative and is labelled as such.

2 The Reconciliation Identity and the Four-Column Proof of Cash

2.1 Two records of one account

Time is measured in days and indexed by integers. Fix one cash account, opened at time 0 with a zero balance on both records (an opening balance is treated as a first receipt). The account has two records.

Definition 2.1 (Cash records)**.** The *book record* is a finite set I^R of ledger receipts and a finite set I^D of ledger disbursements; item i has an amount $a_i > 0$ and a posting date $\tau_i \geq 0$. The *bank record* is a finite set J^R of bank credits and a finite set J^D of bank debits; item j has an amount $b_j > 0$ and a value date $\sigma_j \geq 0$. The book and bank balances at the end of day t are

$$L_t = \sum_{i \in I^R} a_i \mathbf{1}\{\tau_i \leq t\} - \sum_{i \in I^D} a_i \mathbf{1}\{\tau_i \leq t\}, \qquad B_t = \sum_{j \in J^R} b_j \mathbf{1}\{\sigma_j \leq t\} - \sum_{j \in J^D} b_j \mathbf{1}\{\sigma_j \leq t\}.$$

For a period $(s, t]$ with $s < t$, the book receipts and disbursements are $R^{\text{book}} = \sum_{i \in I^R} a_i \mathbf{1}\{s < \tau_i \leq t\}$ and $D^{\text{book}} = \sum_{i \in I^D} a_i \mathbf{1}\{s < \tau_i \leq t\}$, and R^{bank}, D^{bank} are defined in the same way from J^R, J^D and the dates σ_j.

The two roll-forwards are immediate from the definition:

$$B_t = B_s + R^{\text{bank}} - D^{\text{bank}}, \qquad L_t = L_s + R^{\text{book}} - D^{\text{book}}. \tag{1}$$

Each is the conservation law for one record. Neither says anything about the other.

Definition 2.2 (Reconciliation)**.** A *reconciliation* is a pair (μ^R, μ^D), where $\mu^R \subseteq I^R \times J^R$ and $\mu^D \subseteq I^D \times J^D$ are one-to-one partial matchings such that $a_i = b_j$ for every matched pair (i, j). Items in no pair are *unmatched*. The reconciling items at time t are

$$\mathrm{DIT}_t = \sum_{(i,j)\in\mu^R} a_i(\mathbf{1}\{\tau_i \le t\} - \mathbf{1}\{\sigma_j \le t\}), \qquad \mathrm{OC}_t = \sum_{(i,j)\in\mu^D} a_i(\mathbf{1}\{\tau_i \le t\} - \mathbf{1}\{\sigma_j \le t\}).$$

For a set S of items with dates d_ℓ and amounts c_ℓ, write $S(t) = \sum_{\ell\in S} c_\ell \mathbf{1}\{d_\ell \le t\}$ and $S(s,t] = S(t) - S(s)$. Let $U^R_{\text{book}}, U^R_{\text{bank}}, U^D_{\text{book}}, U^D_{\text{bank}}$ denote the four sets of unmatched items.

When the books record every matched item no later than the bank ($\tau_i \le \sigma_j$), these are the textbook objects: DIT_t totals receipts posted by t and credited after t, and OC_t totals disbursements posted by t and debited after t. Items the bank records first, such as a direct credit posted to the ledger days later, enter with a negative sign.

2.2 The stock identity

Theorem 2.3 (Reconciliation identity)**.** *For any reconciliation and any $t \ge 0$,*

$$L_t - B_t = \mathrm{DIT}_t - \mathrm{OC}_t + e_t, \qquad e_t = [U^R_{\text{book}}(t) - U^R_{\text{bank}}(t)] - [U^D_{\text{book}}(t) - U^D_{\text{bank}}(t)]. \quad (2)$$

Proof. Split each of the four sums defining L_t and B_t into matched and unmatched items. On the receipt side, since $a_i = b_j$ on matched pairs and the matching is one to one,

$$\sum_{i\in I^R} a_i \mathbf{1}\{\tau_i \le t\} - \sum_{j\in J^R} b_j \mathbf{1}\{\sigma_j \le t\} = \mathrm{DIT}_t + U^R_{\text{book}}(t) - U^R_{\text{bank}}(t).$$

The same computation on the disbursement side gives $\mathrm{OC}_t + U^D_{\text{book}}(t) - U^D_{\text{bank}}(t)$. Subtracting the second from the first yields (2). □

2.3 The flow identities

Theorem 2.4 (Four-column proof of cash)**.** *For any reconciliation and any period $(s, t]$,*

$$R^{\text{book}} - R^{\text{bank}} = (\mathrm{DIT}_t - \mathrm{DIT}_s) + e^R, \qquad e^R = U^R_{\text{book}}(s,t] - U^R_{\text{bank}}(s,t], \quad (3)$$

$$D^{\text{book}} - D^{\text{bank}} = (\mathrm{OC}_t - \mathrm{OC}_s) + e^D, \qquad e^D = U^D_{\text{book}}(s,t] - U^D_{\text{bank}}(s,t], \quad (4)$$

and the four residuals are linked by

$$e_t - e_s = e^R - e^D. \quad (5)$$

Proof. For a matched receipt pair (i, j), the contribution to $R^{\text{book}} - R^{\text{bank}}$ is $a_i(\mathbf{1}\{s < \tau_i \le t\} - \mathbf{1}\{s < \sigma_j \le t\})$. Since $\mathbf{1}\{s < d \le t\} = \mathbf{1}\{d \le t\} - \mathbf{1}\{d \le s\}$, this equals the pair's contribution to $\mathrm{DIT}_t - \mathrm{DIT}_s$. Unmatched receipts contribute $U^R_{\text{book}}(s,t]$ to R^{book} and unmatched credits contribute $U^R_{\text{bank}}(s,t]$ to R^{bank}, which gives (3); (4) is identical with D in place of R. For (5), take (2) at t and at s and subtract; by the definition of $S(s,t]$ the difference of the bracketed terms is exactly $e^R - e^D$. □

Figure 1 shows the structure. The two rows are the two conservation laws; the four vertical links are the four columns of the proof of cash.

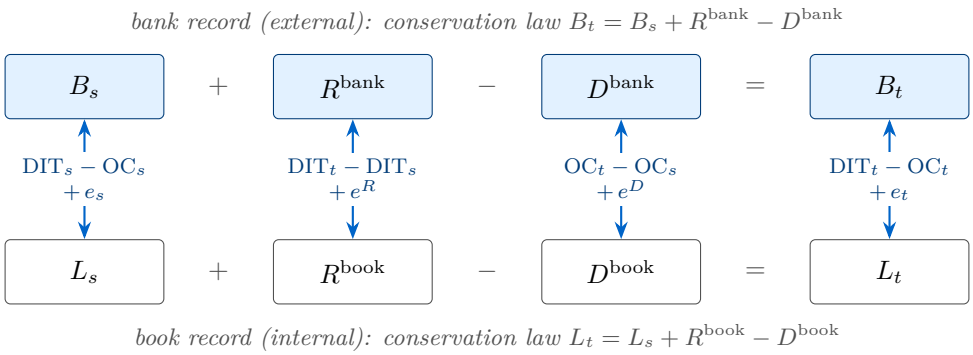


Figure 1: The two roll-forwards and the reconciling items between them. Each vertical link reads "book minus bank". The outer links are the two balance reconciliations of Theorem 2.3; the inner links are the flow columns of Theorem 2.4. The residuals satisfy $e_t - e_s = e^R - e^D$, so the four links carry three independent tests.

2.4 What is definitional and what is content

Definitional. If a preparer simply lists numbers called deposits in transit and outstanding cheques and defines the residuals by (2), (3) and (4), then the link (5) holds by arithmetic, whatever the listed numbers are. A schedule whose columns foot and cross-foot has demonstrated nothing beyond (1).

Content. There are three substantive claims. First, the four bank figures come from a record the company does not write. Second, under the item model the residuals are not free parameters: Theorems 2.3 and 2.4 identify them as the net unmatched amounts, so the assertion "$e^R = 0$" is the testable claim that unmatched ledger receipts and unmatched bank credits of the period have equal totals. Third, and this is the point of the four-column form, the flow columns test two quantities where the balance reconciliations test only their difference.

Corollary 2.5 (Endpoints do not determine flows)**.** *If $e_s = 0$ and $e^R = e^D = 0$ then $e_t = 0$. The converse fails: $e_s = e_t = 0$ holds if and only if $e_s = 0$ and $e^R = e^D$, and the common value of e^R and e^D can be any real number.*

Proof. Both claims are restatements of (5). For the last assertion, take any $x > 0$ and add to a fully matched pair of records one unmatched bank credit of amount x and one unmatched bank debit of amount x, both dated inside $(s, t]$. Then $e^R = e^D = -x$, while e_s and e_t are unchanged. Adding the two items to the book record instead gives $e^R = e^D = +x$. □

Remark 2.6 (The two-line example)**.** A customer payment of 500 is banked and never posted to the ledger; within the same month 500 leaves the bank and is never posted either. Both month-end reconciliations tie, and each flow column shows a variance of −500. Two balance reconciliations are one test of the period; the four-column form is two.

3 Zero Variance Is Not Completeness

A tie-out is a linear functional of the errors in the books: every variance in Section 2 is a signed sum of item amounts. That linearity is the source of both the power of the test and its blind spot.

3.1 The kernel theorem

Let $x \in \mathbb{R}^n$ be an error vector, where x_ℓ is the signed misstatement in the ℓ-th of n independent components (an unmatched item, a line of a schedule, an account-period total). A system of k tie-outs is a matrix $A \in \mathbb{R}^{k \times n}$, and the observed variance vector is $v = Ax$. An error x is *invisible* if $Ax = 0$.

Theorem 3.1 (Kernel theorem)**.** *(a) For a single aggregate variance $v = w^\top x$ with $w \neq 0$, the invisible errors form a subspace of dimension $n - 1$.*

(b) For k tie-outs the invisible errors form a subspace of dimension $n - \operatorname{rank} A \geq n - k$, with equality if and only if the k tie-out equations are linearly independent. If $k < n$ there are invisible errors of every Euclidean norm.

(c) If $A \neq 0$ and x is random with a distribution absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^n$, then $\mathbb{P}(Ax = 0) = 0$.

Proof. (b) The invisible set is $\ker A$, and by the rank-nullity theorem $\dim \ker A = n - \operatorname{rank} A$. Since $\operatorname{rank} A \leq k$, the dimension is at least $n - k$, with equality exactly when the rows of A are linearly independent. If $k < n$ the kernel contains a nonzero vector z, and cz is invisible for every scalar c, which gives every norm. (a) is the case $k = 1$, $\operatorname{rank} A = 1$. (c) Since $A \neq 0$, $\ker A$ is a proper linear subspace of $\mathbb{R}^n$, which has Lebesgue measure zero, hence probability zero under any absolutely continuous law. □

Parts (b) and (c) pull in opposite directions. Part (c) is the reason tie-outs work in practice: an unstructured error is detected almost surely. Part (b) is the reason a zero variance is not a certificate: the errors a tie-out misses are precisely the structured ones, those that offset. A fictitious receipt balanced by a fictitious payment, or a misposting between two covered accounts, lies in the kernel by construction.

The four-column schedule itself illustrates the rank condition. It displays four residuals, yet (5) is a linear dependency among them, so $\operatorname{rank} A \leq 3$ per account-period. For a accounts over P consecutive periods, each ending balance is the next beginning balance, and the independent tests number at most $a(2P + 1)$, however many cells the workbook displays.

There is one setting in which a zero variance is sufficient.

Proposition 3.2 (One-signed errors)**.** *Let $w \in \mathbb{R}^n$ have strictly positive components. If $x \geq 0$ componentwise and $w^\top x = 0$, then $x = 0$.*

Proof. Each term $w_\ell x_\ell$ is nonnegative and the terms sum to zero, so each is zero, and $w_\ell > 0$ forces $x_\ell = 0$. □

A zero variance proves the absence of error only under a model in which all errors push the same way.

3.2 The coverage theorem

The second limit concerns what was never tested. Let $\mathcal{C}$ be the set of account-periods that exist (every bank account of the entity, for every period in scope in which it was open) and let $\mathcal{O} \subseteq \mathcal{C}$ be those for which bank data was obtained.

Definition 3.3 (Coverage)**.** The *coverage ratio* is $\kappa = |\mathcal{O}|/|\mathcal{C}| \in [0, 1]$.

Model the state of the world as $\omega = (\omega_c)_{c\in\mathcal{C}}$, where ω_c is the full pair of records for cell c, and let $x_c(\omega_c) \geq 0$ be the gross cash activity in cell c that is missing from or misstated in the books. The total misstatement is $M(\omega) = \sum_{c\in\mathcal{C}} x_c(\omega_c)$, with observed part $M_{\mathcal{O}} = \sum_{c\in\mathcal{O}} x_c$.

Theorem 3.4 (Coverage theorem)**.** *Assume the admissible states form a product set, so that the state of the unobserved cells is not constrained by the state of the observed cells, and that in each unobserved cell x_c can take every value in $[0, G_c)$ for some $G_c \in (0, \infty]$. Let V be any statistic that is a function of $(\omega_c)_{c\in\mathcal{O}}$ only.*

(a) V is invariant to the state of the unobserved cells. If $\kappa < 1$ and $G_c = \infty$ for some unobserved c, then for every value v attained by V, $\sup\{M(\omega) : V(\omega) = v\} = \infty$.

(b) If $G_c \leq G < \infty$ for all unobserved cells, then $M \leq M_{\mathcal{O}} + (1 - \kappa)\,|\mathcal{C}|\,G$; if $G_c = G$ for all of them, the bound cannot be improved by any function of the observed cells.

Proof. (a) Invariance is immediate, since V does not depend on ω_c for $c \notin \mathcal{O}$. Fix ω with $V(\omega) = v$ and an unobserved cell c with $G_c = \infty$. By the product assumption, for each $m \geq 0$ there is an admissible ω' that agrees with ω on $\mathcal{O}$ and has $x_c(\omega'_c) = m$. Then $V(\omega') = v$ and $M(\omega') \geq m$. (b) There are $|\mathcal{C}| - |\mathcal{O}| = (1 - \kappa)|\mathcal{C}|$ unobserved cells and each contributes less than G. For sharpness, by the product assumption every unobserved cell can be set arbitrarily close to its supremum without changing any observed cell; if all $G_c = G$, $M - M_{\mathcal{O}}$ then approaches $(1 - \kappa)|\mathcal{C}|G$. □

The theorem is elementary and its consequence is not optional. A variance is a statement about $\mathcal{O}$, and nothing licenses extending it to $\mathcal{C}$. The pair (v, κ) is the minimal honest report, and a bound on total misstatement additionally needs the external bound G of part (b), which has to come from other evidence.

Remark 3.5 (Two caveats on κ)**.** First, the product assumption fails for activity that touches a covered account: a transfer from a covered account to an uncovered one leaves a visible one-sided leg (Section 4). The theorem bites on activity that never touches the covered set. Second, the denominator is itself an assertion. If the known set of account-periods $\mathcal{C}_0$ is a strict subset of the true set $\mathcal{C}$, the reported ratio $|\mathcal{O}|/|\mathcal{C}_0|$ overstates the true κ. Completeness of the account list rests on external confirmation [International Auditing and Assurance Standards Board, 2009c], not on the reconciliation.

4 Transfer Neutrality

Diligence usually aggregates the proof of cash across accounts and compares gross receipts with revenue. The aggregate has to be neutral to movements between the entity's own accounts. It is neutral in the net and not in the gross.

Let $\mathcal{K}$ be the set of covered accounts and fix a period $(s, t]$. Every bank line in a covered account is either *external* or one leg of an *internal transfer*. A transfer x has an amount $a_x > 0$, a source account, a destination account, a debit date δ_x in the source and a credit date γ_x in the destination. Call a leg *visible* if its account is in $\mathcal{K}$ and its date is in $(s, t]$. Let G^R and G^D be the gross bank receipts and disbursements summed over $\mathcal{K}$, let R^{ext} and D^{ext} be their external parts, and define

$$T = \sum_{\substack{x:\ \text{both legs} \\ \text{visible}}} a_x, \qquad \rho^R = \sum_{\substack{x:\ \text{credit visible,} \\ \text{debit not}}} a_x, \qquad \rho^D = \sum_{\substack{x:\ \text{debit visible,} \\ \text{credit not}}} a_x.$$

Proposition 4.1 (Transfer neutrality and its failure)**.**

$$G^R = R^{\text{ext}} + T + \rho^R, \qquad G^D = D^{\text{ext}} + T + \rho^D, \tag{6}$$

$$G^R - G^D = R^{\text{ext}} - D^{\text{ext}} + \rho^R - \rho^D. \tag{7}$$

Each residual splits by cause, $\rho^R = \rho^R_{\text{time}} + \rho^R_{\text{cov}}$, where ρ^R_{cov} sums the transfers with a visible credit whose source account is outside $\mathcal{K}$, and ρ^R_{time} sums those whose source is in $\mathcal{K}$ with $\delta_x \notin (s, t]$; the same holds for ρ^D with the roles of the legs exchanged.

Proof. Every visible credit is external or the credit leg of exactly one transfer, and that transfer's debit leg is either visible or not; summing amounts over the three classes gives the first equation in (6). The second is symmetric. Subtracting gives (7), in which T cancels. A debit leg fails to be visible exactly when its account is outside $\mathcal{K}$ or, the account being inside, its date is outside $(s, t]$, which gives the split. □

First, fully visible transfers are neutral in net cash movement and inflate *both* gross columns by T. In an illustrative case with $R^{\text{ext}} = 1{,}200{,}000$ and $T = 300{,}000$, unadjusted gross receipts are 1,500,000, an overstatement of 25 per cent relative to the figure that should be compared with revenue. Second, removing T from both columns leaves external flows plus the residuals ρ^R, ρ^D, which are then the quantities to investigate. Third, the residuals do not cancel in the net unless $\rho^R = \rho^D$. A transfer debited on the last day of the period and credited on the first day of the next adds a_x to ρ^D and nothing to ρ^R.

Corollary 4.2 (Float in the aggregate balance)**.** *Suppose all accounts are covered and opened at time 0 with zero balances. Then the sum of bank balances at t equals cumulative external receipts less cumulative external disbursements, plus*

$$F_t = \sum_x a_x \big(\mathbf{1}\{\gamma_x \le t\} - \mathbf{1}\{\delta_x \le t\}\big).$$

A transfer credited by t and debited after t adds a_x to aggregate reported cash; one debited by t and credited after t subtracts a_x.

Proof. Sum the bank roll-forward (1) from 0 to t over all accounts and separate external lines from transfer legs. Each transfer contributes $+a_x$ once its credit leg has posted and $-a_x$ once its debit leg has posted. □

A positive F_t is cash counted in two places at once, which is why period-end transfers have to be traced through both legs.

5 Matching as Optimisation

Sections 2 to 4 took the matching as given. Producing it is a computational problem with a sharp boundary between easy and hard.

5.1 One-to-one matching is an assignment problem

Let there be n bank lines and m ledger lines on one side of one account. A pair (j, i) is *admissible* if the amounts are equal and the lag $\sigma_j - \tau_i$ lies in a window of length w; an admissible pair has a cost $c_{ji} \in [0, w]$, for instance the absolute lag. A matching is a set of admissible pairs using each line at most once. The natural objective is lexicographic: match as many lines as possible and, among such matchings, minimise the total lag.

Proposition 5.1 (Reduction to assignment)**.** *Choose M with $2M > w \min(n, m)$ and form the square cost matrix of order $n + m$ whose rows are the n bank lines and m dummy rows, whose columns are the m ledger lines and n dummy columns, with entries: c_{ji} for an admissible pair; $2M + 1$ for an inadmissible pair; M for a bank line against a dummy column or a dummy row against a ledger line; and 0 for a dummy row against a dummy column. Then every minimum-cost perfect assignment uses no inadmissible pair, and its real pairs form a maximum-cardinality matching of minimum total cost among maximum-cardinality matchings.*

Proof. Let an assignment use q real pairs. Then $n - q$ bank lines are assigned to dummy columns, $m - q$ ledger lines to dummy rows, and the remaining q dummy rows and q dummy columns to each other. If one real pair (j, i) is inadmissible then $q \geq 1$, so a dummy pair (r, c) exists; replacing (j, i) and (r, c) by (j, c) and (r, i) changes the cost by $2M - (2M + 1) = -1$, so the assignment was not optimal. Hence an optimal assignment corresponds to a matching μ and has cost $\sum_\mu c_{ji} + M(n + m - 2|\mu|)$; conversely every matching arises this way. If $|\mu'| > |\mu|$, then since $0 \leq \sum_{\mu'} c \leq w \min(n, m)$ and $\sum_\mu c \geq 0$,

$$\operatorname{cost}(\mu') - \operatorname{cost}(\mu) \leq w \min(n, m) - 2M\big(|\mu'| - |\mu|\big) < 0,$$

so an optimal μ has maximum cardinality, and among matchings of equal cardinality the cost orders them by total lag. □

The assignment problem is solved exactly in polynomial time by the Hungarian method of Kuhn [1955]. The problem decomposes by amount, since lines with different amounts are never admissible. The objective is a modelling choice and should be stated: the lexicographic rule prefers a distant candidate to an open item, and a reviewer may prefer the opposite.

5.2 One-to-many matching contains subset sum

Deposits are often batched: one bank credit equals the sum of several ledger receipts. The decision problem Batch Match asks, given positive integer amounts $a_1, \ldots, a_m$ in cents and a bank credit of W cents, whether some subset of the ledger lines sums to W.

Proposition 5.2 (Hardness)**.** Batch Match *is NP-complete.*

Proof. It is in NP because a subset is a certificate checkable by one addition pass. Its instances are exactly the instances of the problem that Karp [1972] lists as KNAPSACK (given integers $a_1, \dots, a_m$ and b, decide whether $\sum_i a_i x_i = b$ has a 0-1 solution), now usually called subset sum, which Karp proves NP-complete. The identity map is therefore a reduction in both directions; no property of "amounts in a date window" restricts the instance set. □

The hardness is mostly harmless for ledger data, because amounts are bounded integers. The problem belongs to the knapsack family studied since Dantzig [1957] and has an exact pseudo-polynomial algorithm, which also answers the reviewer's question: is the decomposition unique?

Proposition 5.3 (Exact batch matching with a uniqueness check)**.** *Define* $f_0(0) = 1$, $f_0(v) = 0$ *for* $1 \le v \le W$, *and for* $i = 1, \dots, m$,

$$f_i(v) = \min\{2,\ f_{i-1}(v) + f_{i-1}(v - a_i)\,\mathbf{1}\{v \ge a_i\}\}, \qquad 0 \le v \le W.$$

Then $f_m(W) = \min\{2, N\}$, *where* N *is the number of subsets of the* m *ledger lines that sum to* W. *The table is computed in* $O(mW)$ *additions and comparisons. Thus* $f_m(W) = 0$ *means no batch exists,* 1 *means a unique batch, and* 2 *means the match is ambiguous.*

Proof. Let $N_i(v)$ be the number of subsets of the first i lines with sum v. The subsets counted by $N_i(v)$ split into those that exclude line i, of which there are $N_{i-1}(v)$, and those that include it, of which there are $N_{i-1}(v - a_i)$ when $v \ge a_i$ and none otherwise. So $N_i(v) = N_{i-1}(v) + N_{i-1}(v - a_i)\mathbf{1}\{v \ge a_i\}$, with $N_0(0) = 1$ for the empty set. For nonnegative integers $\min\{2, p+q\} = \min\{2, \min\{2, p\} + \min\{2, q\}\}$, so by induction $f_i(v) = \min\{2, N_i(v)\}$ for all i and v. There are $(m+1)(W+1)$ entries, each computed in constant time. □

A bank credit of $25,000.00 has $W = 2{,}500{,}000$, so a window with $m = 40$ candidates needs 10^8 table entries, which is trivial.

5.3 Coincidental matches

An algorithm that finds a match has not shown that the match is true. Record linkage theory [Fellegi and Sunter, 1969] frames the question correctly: a candidate pair is either a true match or not, and agreement on a field is evidence in proportion to how unlikely the agreement is among non-matches. For an amount field this can be computed.

Proposition 5.4 (Expected coincidental matches)**.** *Let* $p = (p_a)$ *be a probability distribution on amounts, with collision probability* $p_c = \sum_a p_a^2$ *and largest atom* $p_{\max} = \max_a p_a$.

(a) Suppose each of n *bank lines has* m *ledger lines in its window that are not its true counterpart, and that for each such pair the two amounts are independent draws from* p. *Then the expected number of coincidental exact-amount pairs is* nmp_c.

(b) If p *is supported on* K *amounts then* $1/K \le p_c \le p_{\max}$, *with* $p_c = 1/K$ *if and only if* p *is uniform.*

(c) Let $A_1, \dots, A_m$ *be independent with distribution* p *and let* W *be a fixed amount. The expected number of nonempty subsets* S *with* $\sum_{i \in S} A_i = W$ *is at most* $(2^m - 1)p_{\max}$, *and so is the probability that at least one exists.*

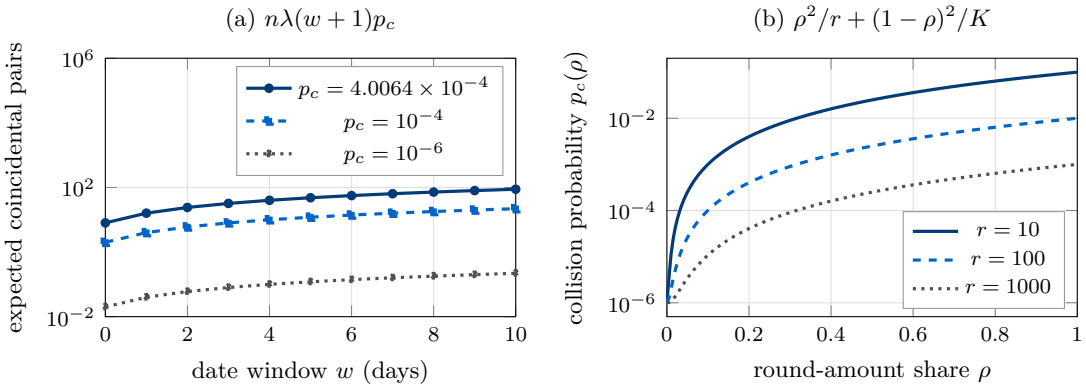


Figure 2: Coincidental exact-amount matches. (a) Equation (8) for illustrative values $n = 1000$ bank lines and $\lambda = 20$ ledger lines per day, for three collision probabilities; the count is linear in $w + 1$ (the vertical axis is logarithmic). (b) Equation (9) with $K = 10^6$: a modest share of round amounts raises p_c by orders of magnitude above the uniform value 10^{-6}.

Proof. (a) For one pair, $\mathbb{P}(\text{equal}) = \sum_a \mathbb{P}(\text{both equal } a) = \sum_a p_a^2$. There are nm pairs and expectation is linear; no independence between different pairs is needed. (b) By the Cauchy-Schwarz inequality, $1 = (\sum_a p_a \cdot 1)^2 \le K \sum_a p_a^2$, with equality if and only if p is constant on its support; and $\sum_a p_a^2 \le p_{\max} \sum_a p_a = p_{\max}$. (c) Fix a nonempty S and some $i \in S$. Conditional on the other amounts, the event requires A_i to equal one specific value, which has probability at most $p_{\max}$. Sum over the $2^m - 1$ nonempty subsets; the last claim is Markov's inequality. □

If ledger lines arrive at a rate of λ per day and the window admits lags $0, 1, \ldots, w$, then $m = \lambda(w + 1)$ and the expected count is

$$\mathbb{E}[\text{coincidental pairs}] = n\,\lambda\,(w+1)\,p_c, \tag{8}$$

linear in the window: widening a date window to clear unmatched items buys matches at a constant marginal price in false ones. Part (c) shows that for batch matching the corresponding bound grows exponentially in the number of candidates.

Real amounts are far from uniform. Digit regularities of the kind described by Benford [1938] and used forensically by Nigrini [2012] are one departure; round amounts are a larger one for this purpose. Suppose a share ρ of lines takes one of r round amounts uniformly and the rest take one of K other amounts uniformly. Then

$$p_c(\rho) = \frac{\rho^2}{r} + \frac{(1-\rho)^2}{K}. \tag{9}$$

With the illustrative values $K = 10^6$, $r = 100$ and $\rho = 0.2$ this gives $p_c = 4 \times 10^{-4} + 6.4 \times 10^{-7} = 4.0064 \times 10^{-4}$, about four hundred times the uniform value 10^{-6}. Figure 2 plots both (8) and (9).

Remark 5.5 (Amount-specific evidence)**.** Under the model of Proposition 5.4, with amounts recorded without error, a true pair agrees on the specific amount a with probability p_a and a non-matching pair does so with probability p_a^2. The likelihood ratio in favour of a true match is $1/p_a$. In the likelihood-ratio framework of Neyman and Pearson [1933], on which the linkage rule of Fellegi and Sunter [1969] is built, agreement on a rare amount such as 4,317.62 is therefore strong evidence, and agreement on 1,000.00 is weak evidence.

6 The Bridge from Cash to Revenue

The proof of cash establishes what was received. Diligence wants to know whether reported revenue is consistent with it. The link is a second pair of roll-forwards.

Over a period, let Rev be revenue recognised, split by how it is recognised: $\mathrm{Rev}^{\mathrm{AR}}$ against a receivable, $\mathrm{Rev}^{\mathrm{cash}}$ against cash at the point of sale, and $\mathrm{Rev}^{\mathrm{DR}}$ by release of deferred revenue. Let $P = P^{\mathrm{AR}} + P^{\mathrm{cash}}$ be pass-through amounts billed to customers and credited to a liability instead of revenue (sales tax is the standard case), split by whether they are added to a receivable or collected directly. Let C^{AR} be cash collected against receivables, Q advance receipts credited to deferred revenue, and X non-cash reductions of gross receivables that do not pass through revenue (write-offs). The roll-forwards of gross receivables and deferred revenue are

$$\Delta\mathrm{AR} = \mathrm{Rev}^{\mathrm{AR}} + P^{\mathrm{AR}} - C^{\mathrm{AR}} - X, \qquad \Delta\mathrm{DR} = Q - \mathrm{Rev}^{\mathrm{DR}}, \tag{10}$$

and cash received from customers is $C = C^{\mathrm{AR}} + \mathrm{Rev}^{\mathrm{cash}} + P^{\mathrm{cash}} + Q$.

Theorem 6.1 (Cash to revenue bridge)**.** *Under* (10)*,*

$$C = \mathrm{Rev} - \Delta\mathrm{AR} + \Delta\mathrm{DR} + P - X. \tag{11}$$

Proof. Solve the first roll-forward for $C^{\mathrm{AR}} = \mathrm{Rev}^{\mathrm{AR}} + P^{\mathrm{AR}} - X - \Delta\mathrm{AR}$ and the second for $Q = \Delta\mathrm{DR} + \mathrm{Rev}^{\mathrm{DR}}$, and substitute:

$$C = (\mathrm{Rev}^{\mathrm{AR}} + P^{\mathrm{AR}} - X - \Delta\mathrm{AR}) + \mathrm{Rev}^{\mathrm{cash}} + P^{\mathrm{cash}} + (\Delta\mathrm{DR} + \mathrm{Rev}^{\mathrm{DR}}).$$

Collecting $\mathrm{Rev} = \mathrm{Rev}^{\mathrm{AR}} + \mathrm{Rev}^{\mathrm{cash}} + \mathrm{Rev}^{\mathrm{DR}}$ and $P = P^{\mathrm{AR}} + P^{\mathrm{cash}}$ gives (11). □

Two details differ from the version usually quoted, $C = \mathrm{Rev} - \Delta\mathrm{AR} + \Delta\mathrm{DR}$. The pass-through term is the amount *billed* in the period, not the amount collected, because billed tax sits inside the receivable and its collection is already captured by $-\Delta\mathrm{AR}$. And write-offs must be subtracted when receivables are measured gross: a write-off reduces the receivable without any cash arriving. As an illustrative check, $\mathrm{Rev} = 500{,}000$, $\Delta\mathrm{AR} = 15{,}000$, $\Delta\mathrm{DR} = 6{,}000$, $P = 40{,}000$ and $X = 2{,}000$ give $C = 529{,}000$.

The bridge variance is $v = C^{\mathrm{bank}} - (\mathrm{Rev} - \Delta\mathrm{AR} + \Delta\mathrm{DR} + P - X)$, where C^{bank} is the customer-receipt total from the proof of cash after transfer elimination and after classifying receipts by source. The classification step is substantive: loan proceeds, owner contributions and asset-sale proceeds are receipts that are not from customers, which is the operating, investing and financing distinction of the cash flow standards [International Accounting Standards Board, 2016, Financial Accounting Standards Board, 2016].

What a tie supports. If $v = 0$, the revenue figure, the two balance movements and the classified bank receipts are mutually consistent. Revenue that was recognised has either been collected or sits in a stated receivable.

What a tie does not support. The bridge is one equation, so Theorem 3.1(a) applies with full force. Overstating Rev and ΔAR by the same amount leaves v unchanged: fictitious credit sales lie in the kernel of the bridge, and only independent evidence on the receivable removes them. More broadly, the tie says nothing about whether the accrual component of earnings will convert to cash, the subject of the accrual literature: Sloan [1996] shows that the accrual and cash components of earnings differ in persistence, Dechow and Dichev [2002] measure accrual quality by how well accruals map into realised cash flows, Richardson et al. [2005] relate persistence to accrual reliability, and Dechow et al. [2010] review the proxies. A tied bridge is compatible with any outcome on these dimensions.

7 An Illustrative Worked Example

Table 1 is a four-column proof of cash for one account and one month. All figures are illustrative. The timing items are deposits in transit of 3,200 at the start and 4,750 at the end, and outstanding cheques of 5,100 and 6,380. The genuine error is the one in Remark 2.6, at a size of 900: a customer receipt banked and never posted, and a bank debit of the same amount never posted.

Table 1: Illustrative four-column proof of cash (one account, one month). The variance row is "per books" less "adjusted bank". Both balance columns tie; both flow columns do not.

	Beginning	Receipts	Disbursements	Ending
Per bank statement	42,300	118,450	109,900	50,850
Deposits in transit, beginning	3,200	−3,200		
Deposits in transit, ending		4,750		4,750
Outstanding cheques, beginning	−5,100		−5,100	
Outstanding cheques, ending			6,380	−6,380
Adjusted bank	40,400	120,000	111,180	49,220
Per books	40,400	119,100	110,280	49,220
Variance	0	−900	−900	0

The arithmetic can be checked line by line. The bank roll-forward is $42{,}300 + 118{,}450 - 109{,}900 = 50{,}850$. The adjusted columns are $42{,}300 + 3{,}200 - 5{,}100 = 40{,}400$; $118{,}450 - 3{,}200 + 4{,}750 = 120{,}000$; $109{,}900 - 5{,}100 + 6{,}380 = 111{,}180$; and $50{,}850 + 4{,}750 - 6{,}380 = 49{,}220$, and they roll forward: $40{,}400 + 120{,}000 - 111{,}180 = 49{,}220$. The book roll-forward is $40{,}400 + 119{,}100 - 110{,}280 = 49{,}220$. In the notation of Theorem 2.4, $R^{\text{book}} - R^{\text{bank}} = 650 = (4{,}750 - 3{,}200) + e^R$ gives $e^R = -900$, and $D^{\text{book}} - D^{\text{bank}} = 380 = (6{,}380 - 5{,}100) + e^D$ gives $e^D = -900$, so $e_t - e_s = e^R - e^D = 0$, as (5) requires.

A reviewer who saw only the two month-end bank reconciliations would see two clean ties. Had the unposted receipt gone into an account outside the covered set, every cell of Table 1 would tie and Theorem 3.4 would be the only relevant statement.

8 Implications for Practice

Report a vector, not a scalar. The output of a proof of cash should be the flow variances per account-period, the coverage ratio κ with its numerator and denominator, the eliminated transfer total T and the residuals ρ^R and ρ^D, and the count of matches that are ambiguous in the sense of Proposition 5.3 or sit on low-evidence amounts in the sense of Remark 5.5.

Keep the computation deterministic and the inputs traceable. Every quantity in this paper is an exact function of ledger rows and bank lines. Finsider accordingly computes a deterministic core from ledger data and keeps any figure extracted by an AI layer separate and labelled as needing verification, so that it does not feed headline metrics until verified. Original source files are kept immutable and reported figures are traced to source rows for the same reason: the residuals of Theorem 2.4 are sums over specific unmatched items, and a variance that cannot be opened into its items cannot be investigated.

Gate on verification, not on a zero. A zero variance is compatible with material error, so it cannot be the release criterion by itself. Finsider holds a company's results behind a release gate until verification passes, with human reviewers in the loop; the reviewer's task is to examine the directions in which the tie-outs are blind. Materiality thresholds [International Auditing and Assurance Standards Board, 2009a] apply to the items so exposed, not to the aggregate variance alone.

9 Falsifiable Predictions

The theorems need no data; their relevance to practice is empirical. The following are predictions about what a future study of real engagements would find, not results.

Prediction 1 (Flow columns find what balance ties miss)**.** Among account-years in which every month-end bank reconciliation ties, a non-trivial share will nonetheless show a material variance in at least one monthly flow column. *Refuted if*, in a sample of at least 200 such account-years, the share with a flow-column variance above the engagement's materiality threshold is below one per cent.

Prediction 2 (Coincidental matches scale linearly in the window)**.** Holding an account's volume fixed, the number of exact-amount candidate pairs that are later judged false will grow in proportion to $w+1$, as (8) states. *Refuted if* a regression of log false-candidate count on $\log(w+1)$ across windows of 0 to 10 days yields a slope whose 95 per cent confidence interval excludes the range $[0.8, 1.2]$.

Prediction 3 (Coverage, not variance, predicts later-found cash misstatement)**.** Conditional on a zero aggregate variance, cash misstatements discovered after the engagement will be concentrated in engagements with $\kappa < 1$. *Refuted if*, among zero-variance engagements, the rate of later-discovered material cash misstatement does not differ between those with $\kappa = 1$ and those with $\kappa < 1$ at the five per cent significance level in a sample with adequate power.

Prediction 4 (A negative prediction: ties do not predict accrual quality)**.** A tied proof of cash together with a tied revenue bridge will not, by themselves, predict the subsequent conversion of period-end receivables to cash. *Refuted if*, controlling for industry and receivable ageing, engagements with both ties show a subsequent write-off rate on period-end receivables

that is lower than engagements with an unresolved bridge variance by a margin that is both statistically significant and larger than one fifth of the mean write-off rate.

10 Limitations

I separate the limits of the procedure from the limits of this analysis.

Limits of the procedure. A proof of cash, even one that ties in every flow column at full coverage with transfers eliminated, does not establish the following. It does not establish revenue recognition: cash received is silent on whether the performance that earns it has occurred. It does not establish ownership or the absence of restrictions on the balance. It does not establish the absence of fraud that operates through the bank record itself: the test treats the bank record as independent, so a falsified statement, or collusion that makes real bank movements mirror fictitious ledger entries (round-tripped funds, for instance), defeats it by design, since such a scheme produces matched pairs. And it does not establish that earnings are sustainable.

Limits of the analysis. The item model of Section 2 uses one-to-one matching with exact amount equality; partial payments, bank fees netted from receipts and foreign currency accounts need additional terms that I have not written out. The kernel theorem is linear; tests with nonlinear structure, such as digit analysis, are outside it. The coverage theorem assumes a product structure that Remark 3.5 qualifies. The count in Proposition 5.4(a) assumes that non-matching amounts are independent draws from a common distribution; recurring payments of a fixed amount violate this. The bridge of Theorem 6.1 assumes three recognition channels and gross receivables; other arrangements (factoring, netting against payables, non-cash consideration) add terms. The predictions of Section 9 are untested, and their thresholds are my judgement of what would count as a clear failure, not estimates.

11 Conclusion

A proof of cash is a conservation law checked against a record the company does not control. Stated at the level of items, its residuals are exactly the net unmatched activity of each period, and its four-column form tests receipts and disbursements separately where balance reconciliations test only their difference. Its limits are equally exact. Tie-outs are linear, so they are blind to an error subspace of dimension at least $n - k$. A variance speaks only for the account-periods it was computed on. Transfers cancel in the net and not in the gross. A match is evidence in proportion to the rarity of the amount. And a tie between cash and revenue is one equation, which fictitious receivables satisfy.

None of this diminishes the procedure. It says what should accompany it: the flow variances, the coverage ratio, the transfer residuals and the ambiguity of the matching. A zero variance is necessary for clean books. The work is in showing what else is true.

References

Frank Benford. The law of anomalous numbers. *Proceedings of the American Philosophical Society*, 78(4):551–572, 1938.

George B. Dantzig. Discrete-variable extremum problems. *Operations Research*, 5(2):266–288, 1957.

Patricia Dechow, Weili Ge, and Catherine Schrand. Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics*, 50(2–3):344–401, 2010.

Patricia M. Dechow and Ilia D. Dichev. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review*, 77(s-1):35–59, 2002.

David P. Ellerman. The mathematics of double entry bookkeeping. *Mathematics Magazine*, 58(4):226–233, 1985.

Ivan P. Fellegi and Alan B. Sunter. A theory for record linkage. *Journal of the American Statistical Association*, 64(328):1183–1210, 1969.

Financial Accounting Standards Board. Accounting standards codification topic 230: Statement of cash flows. FASB, 2016. As amended.

Yuji Ijiri. *The Foundations of Accounting Measurement: A Mathematical, Economic, and Behavioral Inquiry*. Prentice-Hall, Englewood Cliffs, NJ, 1967.

International Accounting Standards Board. IAS 7 statement of cash flows. IFRS Foundation, 2016. As amended.

International Auditing and Assurance Standards Board. International standard on auditing 320: Materiality in planning and performing an audit. IFAC, 2009a.

International Auditing and Assurance Standards Board. International standard on auditing 500: Audit evidence. IFAC, 2009b.

International Auditing and Assurance Standards Board. International standard on auditing 505: External confirmations. IFAC, 2009c.

Richard M. Karp. Reducibility among combinatorial problems. In Raymond E. Miller and James W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, New York, 1972.

Harold W. Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1–2):83–97, 1955.

Jerzy Neyman and Egon S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A*, 231: 289–337, 1933.

Mark J. Nigrini. *Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection*. John Wiley & Sons, Hoboken, NJ, 2012.

Luca Pacioli. *Summa de arithmetica, geometria, proportioni et proportionalita*. Paganino de Paganini, Venice, 1494.

Public Company Accounting Oversight Board. AS 1105: Audit evidence. PCAOB Auditing Standards, 2010.

Scott A. Richardson, Richard G. Sloan, Mark T. Soliman, and Irem Tuna. Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics*, 39(3): 437–485, 2005.

Richard G. Sloan. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review*, 71(3):289–315, 1996.