



# The Evolution of Bank Transaction Data

## Executive summary

Bank transaction data vendors like Yodlee, PLAID, Finicity, DecisionLogic, and MicroBilt have offered products for several years now. Bank transaction data customers, primarily fintech companies, have steadily exploited bank transaction data and have embraced the inherent friction in asking customers to supply bank credentials. Furthermore, the use of bank transaction data adds a new dimension to credit underwriting, facilitating lenders to offer credit to long-underserved segments.

Despite the great promise of the data, bank transaction (BT) applications have generally been limited in scope and depth. BT data is used for identity verification, which involves tying an applicant's name to the underlying bank account and confirming account ownership. Other users have integrated BT data as a node in manual review processes, a practice that limits scale and improved decisions. Some users have gone a step further and have used BT data to create summary statistics on spend types, transaction amounts, debt ratio, and cash flow trends for enhanced underwriting.

The time is right to go deeper; to couple well-known technology with BT data streams, decisively enhance value, and reduce financial exposure. The benefits to BT users, small and large, are improvements in fraud, losses, approval rates, and account management strategies while enabling scalability and depth of insights. Moreover, the availability of robust new data elements implies that high-quality models of credit risk and fraud can be constructed when paired with traditional credit bureau data, internal consumer behavior, application fields, and scores to yield previously unavailable assessments of affordability and incremental risk.



The time is right to go deeper; to couple well-known technology with BT data streams, decisively enhance value, and reduce friction.

## What our solution offers

To address these and other issues, GDS Link and TVD Data Partners have created the BT platform. The platform accepts bank transaction payloads from each major provider and creates a consistent set of data and metadata. These features provide insights into income, expenses, identity, balances, and affordability. BT also consists of advanced machine learning models that assess various financial risks, from first-party fraud to credit risk. These models can be tuned to a specific lending environment in A/B test mode to develop deep insights into loan affordability, payment size, and other metrics that are difficult to measure directly. The BT platform can be used seamlessly with credit bureau data, consumer internal behavior data, credit bureau scores, and—highly recommended—used to build custom scores developed for clients.

## Issues involving bank transaction data

As BT data usage spreads, it is useful to consider two problems in the lending market that it can help answer.

### 1. Score Uncertainty and Score Boosting – an instant boost for credit recovery

On a macroeconomic scale, the lending marketplace increasingly contends with score inflation that can obscure actual marginal risk. Score inflation is created by credit bureaus, aided by a positive economic environment. Its effects are skewed generic scores from FICO and Vantage:



*This means debtors are riskier than their scores indicate because the metrics don't account for the robust economy, skewing perception of borrowers' ability to pay bills on time. When a slowdown comes, there could be a much bigger fallout than expected for lenders and investors. There are around 15 million more consumers with credit scores above 740 today than there were in 2006, and about 15 million fewer consumers with scores below 660, according to Moody's<sup>1</sup>.*

*FICO acknowledges that the credit score alone may not be enough to make informed underwriting decisions, and other factors need to be considered<sup>2</sup>.*



Affordability is best understood by looking backward. Affordability is a revealed phenomenon.

## 2. Affordability – not disposable income, but revealed affordability

All lenders, from prime to nonprime and from online to brick and mortar, would benefit from assessing important information related to their customers' financial status. Latent aspects, such as affordability, are difficult to deduce from credit bureau data and are subject to noisy or misleading answers from applicants. Access to up-to-date bank transactions is an excellent way to get greater insight into an applicant's ability to pay; access to hundreds of derived features would add even more value at the margin.

Applicants and existing customers may not always know when or whether they can afford a loan or a refinance. After all, many loans are taken when a customer is already short of cash. The problem is exacerbated when considering that many potential customers (and often the most profitable ones) have insufficient bureau data for an informed underwriting decision. To account for these issues, regulatory pressure has developed to induce creditors to better understand affordability.

Affordability is best understood by looking backward. Affordability is a revealed phenomenon. Conceptual problems can be avoided altogether by considering affordability as the latent entity that it is. For example, disposable income is often assessed using circumstantial data and is a primary indicator of affordability. Yet disposable income is never truly observed because it's difficult to determine if retail spending (for example) is

for tires or groceries. Moreover, not all bills are labeled, so utilization of funds can be hard to classify.

BT data, which provides deep insight into revealed affordability, allows lenders to set payments on dates best suited to borrower, identifies remaining balance at time periods, and indicates trends about outflows and overdrafts.

Snapshot statistics do not offer much apart from describing current balance. An important decisioning component, categorization, is often wrong or missing. Existing solutions depend on transaction categorization and every BT provider takes a different approach to the issue. However, the categorization challenge across the industry is clear:



Up to 20% of transactions are not categorized.

- Up to 20% of transactions are not categorized
- Up to 5% of inflows are not categorized or are misclassified
- Up to 15% of outflows are not categorized or are misclassified
- Many things end up in 'Other'

We have observed that the quality of the data is not equal and such quality varies with provider. For example, many of the merchant names or lenders are written in different ways in the same file depending on transaction provider.

These variations come in many forms of the same name (ex.- Walmart, Wal-Mart, WalMart).

The above issues have led to BT data providers grouping merchants' names or categories in 'Other' income or spending.

When using GDS Link BT attributes, it will not be necessary to spend time categorizing transactions. Categorization requires deployment of precious analytic resources plus long periods of research and testing to validate and make sure the names are always updated. By using GDS Link's well-managed BT data, the focus can rest on trended velocity, periodicity, and repetitiveness, rather than on brittle, time-consuming categorization.



By using GDS Link's well-managed BT data, the focus can rest on trended velocity, periodicity, and repetitiveness, rather than on brittle, time-consuming categorization.



# GDS Link Bank Transaction Platform

## The Solution offers:

- Normalized view of bank transaction data regardless of provider
- 4,500+ enhanced bank transaction features that are customizable to the client
- Improved categorization
- Flexible levels of KYC/KYB
- Refreshed flags and triggers
- Model monitoring package
- Re-tuning and deployment of models as a service
- Access to dashboard statistics

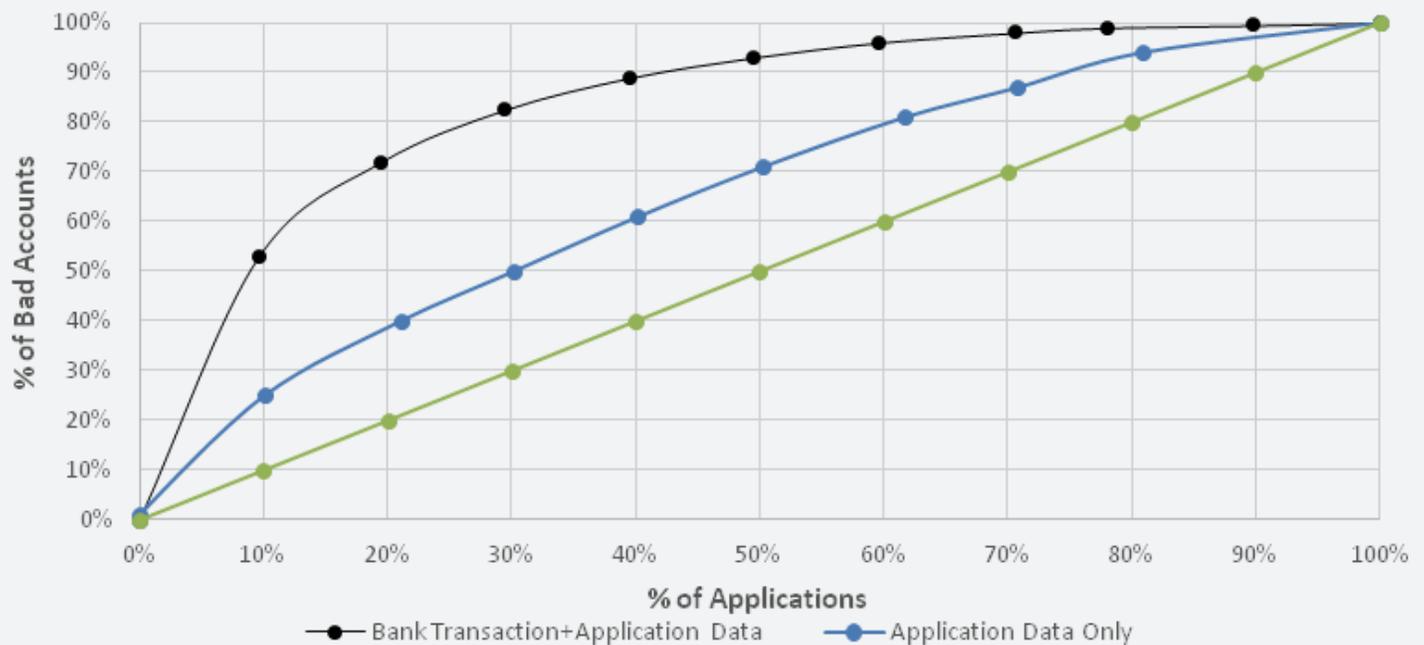
GDS Link, in partnership with TVD Data Partners, has created a suite of modules based on BT data that provides an array of trended attributes and flags. These can be accessed on the GDS Link platform via an API. In addition to trended variables, data features are based on several themes: Inflows, Outflows, Overdraft, Contactability, Balance, and Account. The data solution is developed and maintained by data and predictive analytics professionals with over 30 combined years of experience in data science and technology and their application of BT in the financial sector. The platform's benefits are equally compatible to prime, non-prime, deep sub-prime, and thin-file segments for consumer and small lending businesses.

The GDS Link Bank Transaction Platform allows real-time application flow or batch access. Its cloud-based architecture can accept data payloads from any bank transaction data provider. GDS Link will have credit and fraud scores available for use. Concurrently, client-developed machine learning or multivariate models can be hosted and scored in either environment. The platform supports rapid re-tuning and redeployment of models with simple business rules is supported to drive benefits sooner. The API-driven GDS Link technology means that all native languages are supported.

## GDS Link efficiency & data enriching

The critical aspect of BT data is its richness. Even a month of bank transactions can hold over 100 individual transactions, ranging from debits to credits, and including checking and savings account utilization. Payroll transactions, whether in paper check or direct deposit form, represent inflows, and their amount and frequency are the basis of affordability assessments. Debits in the form of paper checks, ACH transactions, and ATM usage sum to provide a unique profile of a potential customer. Deposits and running balances can be weighed against debits, insufficient funds notices, and zero/negative balance instances to make effective assessments of ability to pay from the account or accounts in question. BT data is highly atomic, hence deep insights into ability and willingness to pay requires

Exhibit A: Bank Transaction Data incorporated with Application data Vs Application Data only



a platform that can quickly and accurately derive meaning from such data.

GDS Link is constantly evolving our solution by building a suite of enhanced features and summary statistics to “dig deeper” into the customer’s profile, around first party fraud and likelihood of fraud. We have found that combining this suite with deep learning, machine learning, and traditional approaches yields a substantial lift in accuracy and efficiency.



We have found that combining this suite with deep learning, machine learning, and traditional approaches yields a substantial lift in accuracy and efficiency.

## Solutions for BT data granularity, incomplete profiling & identity theft

Managing high frequency, atomic data can be difficult and cumbersome. Even sophisticated data aggregations and lending platforms generally accept data flows from API feeds and make limited calculations from the granular data. What is required is a detailed, replicable process that accepts the atomic data and creates a sufficiently rich set of individual features, which can be used in lending decisions or in credit scoring. For example, a data process should be able to identify all payroll activity and manipulate it to provide insight into the amount and consistency of income. Likewise, certain atomic debits can be summed to provide useful insights into spending amounts, trends, categories, and outflow in relation to income. The metadata potentially sourced from BT data requires a scalable and replicable process that functions between APIs accessing bank data, and underwriting criteria using data and scores.

GDS Link provides a more evolved set of options to determine the KYC or KYB aspect of onboarding a customer.

**We have found that using name matching by itself creates an incomplete view. A triage approach is more stable, flexible, and reduces risk.**

- a. **Application Transactions:** Does the applicant name match the transaction's name? Deepen the match logic against payroll or money transfer.
- b. **Account Transactions:** Does the account holder name match the transaction's name? Ongoing verification against payroll or money transfers over time.

## Control over identity account takeover

- a. Above, if a fraudster takes over a bank account they can change the account holder name. However, they cannot change the incoming payroll name which reflects the real owner's name (e.g. ADP payroll smith 4119). We have defined logic that would flag name mismatches.
- b. **Changes in contact history.** Snapshot identity metrics from bank transaction providers only provide users with snapshots at that particular time. We focus on changes to BII/PII elements over time. Does the match rate to app BII/PII info or account BII/PII change over time (30 vs 60 vs 90 day windows)?
- c. **Why do we check phone, address, and e-mail?**  
By checking the location-based matching to see if spending location has changed for fraud purposes. Many transactions list a city, state or city, state, zip or a phone number (area code). This means we should be looking for those changes over a 30, 60, 90 day window.
  - A dashboard graph network, for directional spending behavior, and a dashboard of spending by week and location on a map, can be shared with the customer or used internally.

*(continued)*

## Control over identity account takeover continued

- d. Disposable income problems: Only a few providers supply users with some information summary; the current snapshot and such a summary will not offer users all spending/income sources, hence the resulting disposable income number will always be incorrect.
- Is a Walmart expense for groceries OR over-the-counter medicines OR toiletries OR tires?
  - Do transfers between accounts and people hide the reasoning disposable income requires?
  - Banking institutions change transaction categorization on impulse, which impacts categorization and aggregation by BT providers.
  - GDS Link & TVD Data Partners use revealed affordability in the form of velocity changes, inflow and outflow weekly changes, and trended metrics to determine changes in consumption and income, which in turn reflect external sources of spending and income not directly provided in the account.

The nuanced capabilities of our BT data platform add up to transform how firms can leverage transaction information in everyday operations. The results are evident in both consumer- and commercial-focused operations.

## BT data in consumer services



BT information can help banks move on from unwieldy, inconvenient fraud prevention methods.

Leveraging BT information gives lenders an opportunity to improve customers' experiences by providing more sophisticated product offerings that are tailored to specific consumer needs. For example, if analysis of transaction data shows that a customer can afford to spend more on housing, a bank may choose to send that individual or household marketing materials about mortgages, either for primary residents or secondary housing. This is just one option, of course, but the granular details offered by transaction data provide a real opportunity for personalization within banking services.

Beyond that, BT data offers the kind of visibility into risk that lenders need to expand their products to new audiences. Whether you're hoping to reach invisible consumers without a credit history or focus on services for individuals with bad credit, the cash flow analysis provided by the BT platform gives you the visibility you need.

The fraud prevention advantages of BT data are particularly beneficial in the consumer lending sector, where the ability to identify trends in customer spending and flag potentially fraudulent expenses becomes invaluable. BT information can help banks move on from unwieldy, inconvenient fraud prevention methods and implement more strategic, customer-friendly options that provide vital protection with minimal disruption.

## BT data for commercial banking

The alternative lending boom has hit many banks and credit unions hard. Online lenders have fueled expectations for faster loan processing across all audiences, but the effects have been especially evident in small business lending. Many banks struggle to ensure profitable small business lending as the relatively small size of loans combines with the risk of borrowers in that sector to create high costs and limited returns.

BT data provides more visibility into business cash flow, letting banks perform deeper analysis on what borrowers can actually afford, all in an efficient timeline. The deep data analysis is instrumental in accelerating the lending process while reaching a wider range of audiences.

On top of lending, the fraud prevention and transaction-level visibility can help banks provide more varied and robust business banking services.



The deep data analysis is instrumental in accelerating the lending process while reaching a wider range of audiences.

## Why we like the evolved state of bank transaction data

The uniqueness of the variables in BT data analysis is based on the ability to capture trends versus snapshots through improved categorization and classification. The power of the information once fed into models enhances the credit and internal Masterfile customer data. More importantly, it captures specific behaviors and insights into customers' ability to use, pay, and consume loan products.

This data is unique in that it can feed personal finance management tools to help consumers manage spending and track where cash comes in and goes out in real time. Using the data to communicate debt to income, spend categories, and cash frequency can be a valuable feedback loop to engage customers.



It captures specific behaviors and insights into customers' ability to use, pay, and consume loan products.

GDS Link has built a scalable, configurable data solution which offers bank transaction insights end-to-end across customer lifecycles. These tools can be applied to address account verification, fraud assessments, credit risk, and affordability. The GDS Link solution is accompanied by a robust, tested decisioning engine and available consulting services to enhance underwriting, digital marketing, and sophisticated acquisition workflows.

## **About GDS Link**

GDS Link is a global leader in credit risk management, providing tailored software solutions, analytical and consulting services. Our customer-centric risk management and process automation platforms are designed for the modern lender in their pursuit to capitalize on the entire credit lifecycle. By providing a personal, consultative approach and leveraging our own industry-leading knowledge and expertise, GDS Link's solutions and services deliver exceptional value and proven results.

[gdslink.com](https://gdslink.com) | [info@gdslink.com](mailto:info@gdslink.com) |



**GDS LINK**